

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 15:03:04 ; Search time 29 seconds
(without alignments)
77.327 Million cell updates/sec

Title: US-10-016-768a-2

Perfect score: 273
Sequence: 1 KOPRRKRGROYDHEIMEE.....KAOIGVPHSTLEYKVER 53

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	59	21.6	191	2 US-08-286-819A-21	Sequence 21, Appl
2	59	21.6	191	3 US-08-980-357-21	Sequence 21, Appl
3	57	20.9	102	4 US-09-732-210-1293	Sequence 1293, Ap
4	56	20.5	696	4 US-09-107-532A-6133	Sequence 6133, Ap
5	55.5	20.3	104	4 US-09-732-210-1272	Sequence 1272, Ap
6	54	19.8	625	4 US-09-252-991A-33087	Sequence 33087, A
7	53.5	19.6	187	4 US-09-328-332-5117	Sequence 5117, A
8	53.5	19.6	438	4 US-09-134-001C-3634	Sequence 3634, Ap
9	53	19.4	137	4 US-09-107-532A-4469	Sequence 4469, Ap
10	53	19.4	227	4 US-09-308-003-48	Sequence 48, Appl
11	53	19.4	390	4 US-09-308-003-12	Sequence 12, Appl
12	52.5	19.2	389	4 US-09-252-991A-24471	Sequence 24471, A
13	52.5	19.2	810	4 US-09-540-824-25	Sequence 25, Appl
14	52	19.0	552	4 US-08-974-546-1	Sequence 1, Appl
15	52	19.0	552	4 US-09-446-301A-4	Sequence 4, Appl
16	52	19.0	552	4 US-09-099-932-4	Sequence 4, Appl
17	52	19.0	560	4 US-09-446-301A-50	Sequence 50, Appl
18	52	19.0	663	1 US-07-912-015-2	Sequence 2, Appl
19	51.5	18.9	349	3 US-08-462-467B-18	Sequence 18, Appl
20	51.5	18.9	554	3 US-08-462-467B-22	Sequence 22, Appl
21	51.5	18.9	1399	3 US-08-462-467B-14	Sequence 14, Appl
22	51	18.7	101	4 US-09-732-210-1268	Sequence 1268, Ap
23	51	18.7	218	4 US-09-107-532A-5919	Sequence 5919, Ap
24	51	18.7	222	4 US-09-252-991A-20670	Sequence 20670, A
25	51	18.7	425	3 US-08-462-467B-16	Sequence 16, Appl
26	51	18.7	649	3 US-08-462-467B-20	Sequence 20, Appl
27	51	18.7	1495	3 US-08-462-467B-12	Sequence 12, Appl

28	50.5	18.5	113	4 US-09-107-532A-6213	Sequence 6213, Ap
29	50.5	18.5	521	4 US-09-252-991A-30932	Sequence 30932, A
30	50	18.3	60	3 US-08-928-213B-184	Sequence 184, Appl
31	50	18.3	100	2 US-08-160-524A-12	Sequence 12, Appl
32	50	18.3	278	3 US-08-821-278A-18	Sequence 18, Appl
33	50	18.3	346	2 US-08-613-220B-4	Sequence 4, Appl1
34	50	18.3	398	4 US-09-252-991A-21127	Sequence 21127, A
35	50	18.3	519	4 US-09-107-532A-7161	Sequence 7161, Ap
36	50	18.3	802	3 US-09-081-345-18	Sequence 18, Appl
37	49.5	18.1	697	4 US-09-252-991A-27283	Sequence 27283, A
38	49	17.9	1209	4 US-08-252-991A-25844	Sequence 25844, A
39	49	17.9	2584	3 US-08-936-135-4	Sequence 4, Appl1
40	48.5	17.8	102	4 US-09-732-210-1283	Sequence 1283, Appl
41	48.5	17.8	104	4 US-09-732-210-144	Sequence 144, Appl
42	48.5	17.8	352	4 US-09-996-243-505	Sequence 505, App
43	48.5	17.8	361	4 US-09-198-452A-812	Sequence 812, App
44	48.5	17.8	365	2 US-08-979-424-3	Sequence 3, Appl1
45	48.5	17.8	365	3 US-08-928-383B-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-286-819A-21
: Sequence 21, Application US/08286819A
: Patent No. 5871910
: GENERAL INFORMATION:
: APPLICANT: ARTHUR, MICHEL
: APPLICANT: DUKTA-MALEN, SYLVIE
: APPLICANT: MOLINAS, CATHERINE
: APPLICANT: COUVAILIN, PATRICE
: TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
: TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
: TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
: NUMBER OF SEQUENCES: 54
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: STREET: 1755 S. Jefferson Davis Highway, Suite 400
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/286,819A
: FILING DATE: 05-AUG-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/174,682
: FILING DATE: 28-DEC-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/917,146
: FILING DATE: 10-AUG-1992
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/FR/91/00855
: FILING DATE: 29-OCT-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 9013579
: FILING DATE: 31-OCT-1990
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Oblon, No. 5871910man F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 660-060-0 PCT

```

      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (703) 413-3000
      TELEFAX: (703) 413-2220
      TELEX: 248855 OPAT-UR
      INFORMATION FOR SEQ ID NO: 21:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 191 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-286-819A-21

Query Match      21.6%; Score 59; DB 2; Length 191;
Best Local Similarity 29.6%; Pred. No. 0.71;
Matches 16; Conservative 13; Mismatches 23; Indels 2; Gaps 1;

Db      135 KKEGFKGRLLKHKHKNAGMNVAVKLYKEGNTVNOICEITVYRSALYKRLSE 188
      1 KQPKKAGRRQY--DHEIMEEAIAMVSGKMSVSKAGCIGVPHSTLEYKYKE 52
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 2
; Sequence 21, Application US/08980357
; Patent No. 6013508
; GENERAL INFORMATION:
; APPLICANT: ARTHUR, MICHEL
; APPLICANT: DUKTA-MALEN, SYLVIE
; APPLICANT: MOLINAS, CATHERINE
; APPLICANT: COURVALIN, PATRICE
; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPETIDES, IN PARTICULAR
; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/980,357
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,819
; FILING DATE: 05-AUG-1994
; APPLICATION NUMBER: US 08/174,682
; FILING DATE: 28-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/917,146
; FILING DATE: 10-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR/91/00855
; FILING DATE: 29-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9013579
; FILING DATE: 31-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Obion, No. 6013508man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-060-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220

```

```

TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-357-21

Query Match
Best Local Similarity 21.6%; Score 59; DB 3; Length 191;
Matches 16; Conservative 13; Mismatches 23; Indels 2; Gaps 1;

QY
1 KQPKRRRYQY--DHEIMEEAIAMWMSGMSVSKAGIYGVPHSTLEKVK 52
: : : : : : : : : : : : : : : : : : : : : : : : :
DB 135 KKEGFKGRLLKRYHKNHAGMNYAVKLYKEGMVMTQICEITNVSRASTYRLSE 188

RESULT 3
US-09-732-210-1293
Sequence 1293, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Miltanek, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-Fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 1293
LENGTH: 102
TYPE: PRT
ORGANISM: Thermotoga maritima
US-09-732-210-1293

Query Match
Best Local Similarity 20.9%; Score 57; DB 4; Length 102;
Matches 15; Conservative 14; Mismatches 17; Indels 10; Gaps 2;

QY
4 RKKRRRYQYDHEIMEEAIAMWMSGMSVSKAG-----IYGVPHSTLEK 49
: : : : : : : : : : : : : : : : : : : : : : : : :
DB 4 QKIRIKLAYDHELDESARKRIVEAKSTNSKVGPIPLPTERTLYCLVLRSPMKH 59

RESULT 4
US-09-107-532A-6133
Sequence 6133, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
CORRESPONDENCE ADDRESS: ENTEROCOCCUS FACIUM FOR DIAGNOSTICS AND THERAPEUTICS
ADDRESS: GENOME THERAPEUTICS CORPORATION
ADDRESS: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneko
REGISTRATION NUMBER: 40,489
TELECOMMUNICATION INFORMATION:
REFERENCE/DOCKET NUMBER: GTC-012
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6133:
SEQUENCE CHARACTERISTICS:
LENGTH: 696 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...696
SEQUENCE DESCRIPTION: SEQ ID NO: 6133:
US-09-107-532A-6133

Query Match
Best Local Similarity 20.5%; Score 56; DB 4; Length 696;
Matches 18; Conservative 8; Mismatches 18; Indels 12; Gaps 3;

QY 3 PRKKRGYROYDHEIMEAIAVMWSGKMSVSKAOGITGVPHSTLEYKRV-----KER 53
DB 372 PTEENRHYEKLSE-ERKTY-----OMIVSRFLGLAOPHKYSOTFVTEFGKER 420

RESULT 5
US-09-732-210-1272
Sequence 1272, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mitlanck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 1272
LENGTH: 104
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-732-210-1272

Query Match
Best Local Similarity 20.3%; Score 55.5; DB 4; Length 104;
Matches 14; Conservative 12; Mismatches 20; Indels 1; Gaps 1;

QY 5 KRGRYROYDHEIMEEA-IAVMWSGKMSVSKAOGITGVPHSTLEYKV 50
DB 3 KIRKLKAYDHRVLDRSVAIVAEVAKRSGSEIRGPILPKNKRTYV 49
```

```

RESULT 6
US-09-252-991A-33087
Sequence 33087, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 33087
LENGTH: 625
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33087

Query Match
Best Local Similarity 19.8%; Score 54; DB 4; Length 625;
Matches 13; Conservative 7; Mismatches 12; Indels 4; Gaps 1;

QY 6 KRGRYROYDHEIMEAIAVMWSGKMSVSKAOG 37
DB 90 ERGHRRLGIAFQHEVAEAREAPVGRLEQLAEG 125

RESULT 7
US-09-328-352-5117
Sequence 5117, Application US/09328352
Patent No. 6562858
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTE
FILE REFERENCE: GTC99-038A
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 5117
LENGTH: 187
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-5117

Query Match
Best Local Similarity 19.6%; Score 53.5; DB 4; Length 187;
Matches 15; Conservative 12; Mismatches 16; Indels 5; Gaps 2;

QY 10 YROYDHEIMEAIAVMWSGKMSV-SKAOGIV---GVPHSTLEYKRV 52
DB 64 YQQRNVEIKIEVINLFRISDSIHKKLKGVDLTGKNSDEPKTYE 111

RESULT 8
US-09-134-001C-3634
Sequence 3634, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
```

```

: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 3634
: LENGTH: 438
: TYPE: PR1
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3634

```

Query Match	19.6%	Score 53.5	DB 4	Length 438
Best Local Similarity	38.2%	Pred. No. 14		
Matches 13, Conservative	5	Mismatches 5	Indels 11	Gaps 1

```

Qy      4  RKKRGRRQYDHD-----ELMEEAALAVM 26
          ||::|||||::          | | :|||
Db      129 RKKRGRRQFNQGVGAIGAPNSIDAEILAVM 162

```

RESULT 9
US-09-107-532A-4469
: Sequence 4469, Application US/09107532A
: Patent No. 6583275
: GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:

ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Waltham
STATE: Massachusetts
COUNTRY: USA

```

;
;      COMPUTER READABLE FORM:
;      ZIP: 02354
;
;      MEDIUM TYPE: CD-ROM ISO9660
;      COMPUTER: PC
;      OPERATING SYSTEM: <Unknown>
;      COMMAND: accit

```

```

;
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/107,532A
;   FILING DATE: 30-Jun-1998
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 60/005,500

```

APPLCATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLCATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: [REDACTED]

NAME: Arniello, Pamela Denene
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
FAX: (781)893-8332

TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 4469:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 137 amino acids
 TYPE: amino acid
 ORGANISM: Homo sapiens

```

;          TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          HYPOTHETICAL: YES
;          ORIGINAL SOURCE:..
;          ORGANISM: Enterococcus faecium

```

```

;      FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION: (B) LOCATION 1...137
;      SEQUENCE DESCRIPTION: SEQ ID NO: 4469:
US-09-107-532A-4469

```

US-09-107-532A-4469				
Query Match	19.4%;	Score 53;	DB 4;	Length 137;
Best Local Similarity	21.4%;	Pred. No. 3.5;		
Matches	12;	Conservative	16;	Mismatches 18;
				Indels 10;
				Gaps 1

```

0Y      4 RKKRGKROYDHEIMEEALMV-----MSGKMSVSKAQGYCVPHSTLEK 45
      :| | | :| | :| | :| | :| | :| | :| | :| | :| | :| |

```

DB 66 QIRIRLKAYEHRILDQADKIVETAKRTGASVSGPIPLPTERSLYTIIRATHKKYK 121

RESULT 10
US-09-308-003-48
; Sequence 48, Application US/09308003
; Data# No. 6326170

; APPLICANT: Burnham, Martin K. R.
 ; APPLICANT: Lonetto, Michael A.
 ; APPLICANT: Warren, Patrick V.
 ; TITLE OF INVENTION: NOVEL DEKRAVOTIC POLYMERIC ESTERS

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

? EARLIER APPLICATION NUMBER: 86/0958 /10
? EARLIER FILING DATE: 1997-09-12
? NUMBER OF SEQ ID NOS: 52
? SOFTWARE: FastSeq for Windows Version 3.0.
? SEQ ID NO 48
? LENGTH: 227

```

```

; LENGTH: 22/
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-308-003-48

```

Query Match	19.4%	Score 53	DB 4	Length 227
Best Local Similarity	28.3%	Prod. NC	6.8	
Matches 13, Conservative	10	Mismatches	15	Indels 8; Gaps 1;

```

QY      12 QYDH-----EIMEEAIAMVMSGKMSVSKAQGIYGVPHSTLEK 49
      :::: |::: |::: |::: |::: |::: |::: |::: |::: |:::
Db      49 KYDYARSGNPBRELLEKRLARLEQKFAFAFGIAAISAVLLTEK 94

```

RESULT 11
US-09-308-003-12
; Sequence 12, Application US/09308003
; Date of No. 6326170

Patent No. 6326170
GENERAL INFORMATION:
APPLICANT: Burnham, Martin K. R.
APPLICANT: Lometto, Michael A.
APPLICANT: Warren, Patrick V.
TITLE OR INVENTION, NATURE OF INVENTION, POLYMERIZATION

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

1 EARLIER APPLICATION NUMBER: 60/058,710
2
3 EARLIER FILING DATE: 1997-09-12
4
5 NUMBER OF SEQ ID NOS: 52
6
7 SOFTWARE: FASTSEQ for Windows Version 3.0.
8
9 SEQ ID NO 12
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
102
```

```

; LENGTH: 390
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-308-003-12

```

Query Match	19.4%	Score	53	DB	4	Length	390
Best Local Similarity	28.3%	Pred. NC	14				
Matches	13	Conservative	10	Mismatches	15	Indels	8
						Gaps	1

```
QY      12 QYDH-----EIMEAIAAMVMSGKMSYSKAOGIYGVPHPSTLEFK 495
       :|: |::|| |: |:: ||: |:: |:: |:: |:: |:: |:: |:: |::
Db     42 KYDYARSGNPNNELLEKLRLAQGKFAPAFASGIAAISAVLLTFK 87
```

RESULT 12
US-09-252-991A-24471
; Sequence 24471, Application US/09252991A

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al

```

```

: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 24471
: LENGTH: 389
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24471

Query Match      19.2%; Score 52.5; DB 4; Length 389;
Best Local Similarity 22.8%; Pred. NO. 16;
Matches 13; Conservative 15; Mismatches 16; Indels 13; Gaps 2;

OY      8 GRVROYDHEIMEAIA-----MVMSGKMSVSKAGICVPHSTLEKVKER 53
DB      217 GEFOGAHEADRAVAGQRLLEVEDYVAGAGIASDQHY--PHSRIGLRLOQR 271.

RESULT 13
US-09-540-824-25
: Sequence 25, Application US/09540824
: Patent No. 6383753
: GENERAL INFORMATION:
: APPLICANT: Thiele, Dennis
: TITLE OF INVENTION: No. 6383753el Yeast and Mammalian Regulators of Cell Proliferat
: FILE REFERENCE: UM-04286
: CURRENT APPLICATION NUMBER: US/09/540,824
: CURRENT FILING DATE: 2000-03-31
: NUMBER OF SEQ ID NOS: 28
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 25
: LENGTH: 810
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae
US-09-540-824-25

Query Match      19.2%; Score 52.5; DB 4; Length 810;
Best Local Similarity 36.8%; Pred. NO. 42;
Matches 14; Conservative 7; Mismatches 12; Indels 5; Gaps 1;

OY      1 KOPRRKRGROYDHEIMEAIAVMSGKMSVSKAGI 38
DB      199 QOPORRKT-----AEVWKEVIAKSKFYKOROKRAGI 231

RESULT 14
US-08-974-546-1
: Sequence 1, Application US/08974546
: Patent No. 5645287
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Lal, Preeti
: APPLICANT: Corley, Neil C.
: APPLICANT: Shah, Purvi
: TITLE OF INVENTION: TWO HUMAN HEAT SHOCK PROTEIN HOMOLOGS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Dr.
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
```

```

: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/974,546
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0428
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 348 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: BRAITUT21
: CLONE: 2525691
US-08-974-546-1
```

```

Query Match      19.0%; Score 52; DB 2; Length 348;
Best Local Similarity 39.0%; Pred. NO. 17;
Matches 16; Conservative 4; Mismatches 19; Indels 2; Gaps 1;
```

```
OY      5 KKRGRYROYDHEIMEAIAVMSGKMSVSKAGICVPHSTL 45
DB      59 KRGGLDYDGEGLKTGGC--TSGGSGSFHYTFGDPHAT 97
```

```

RESULT 15
US-09-446-301A-4
: Sequence 4, Application US/09446301A
: Patent No. 6506893
: GENERAL INFORMATION:
: APPLICANT: EL SOLH, NEVINE
: TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
: FILE REFERENCE: 03715-0059
: CURRENT APPLICATION NUMBER: US/09/446,301A
: CURRENT FILING DATE: 1999-12-20
: NUMBER OF SEQ ID NOS: 51
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 4
: LENGTH: 552
: TYPE: PRT
: ORGANISM: Staphylococcus sp.
US-09-446-301A-4
```

```

Query Match      19.0%; Score 52; DB 4; Length 552;
Best Local Similarity 31.1%; Pred. NO. 30;
Matches 14; Conservative 11; Mismatches 12; Indels 8; Gaps 2;
```

```
OY      4 RKRGRYROYD--HEIMEAIAVMSGKMSVSKAGICVPHSTL 46
DB      193 RKQDEYKYNKRKOLEDAI-----KLKENKAGMTPPSKTM 231
```

```

Search completed: August 4, 2003, 15:08:06
Job time : 30 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 14:55:14 ; Search time 26 Seconds
(without alignments)
95.862 Million cell updates/sec

Title: US-10-016-768a-2
Perfect score: 273
Sequence: 1 KQPRKKRGYRQYDHEIMEE.....KAQIGVPHSTLEYKVKER 53

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69.5	25.5	753	1	TKR_DROME
2	63	23.1	977	1	BAB1_DROME
3	61	22.3	352	1	YFDF_ECOLI
4	59	21.6	191	1	TNR6_ENTFC
5	59	21.6	1067	1	BAB2_DROME
6	58	21.2	342	1	REC1_MYXXA
7	58	21.2	397	1	C1ST_STNY3
8	57	20.9	102	1	RS10_THEMA
9	56.5	20.7	364	1	SYM_MERTH
10	56	20.5	102	1	RS10_LACIA
11	56	20.5	102	1	RS10_LISMO
12	55.5	20.3	104	1	RS10_HELPY
13	55.5	20.3	169	1	Y769_METJA
14	55	20.1	236	1	Y416_CHLTR
15	55	20.1	420	1	SYH_STIAM
16	55	20.1	765	1	PURL_ARCFU
17	55	20.1	809	1	QUTR_ACICA
18	55	20.1	882	1	ALAS_PSEOL
19	54.5	20.0	2607	1	BACB_BACLI
20	54	19.8	352	1	RECA_BORPE
21	54	19.8	764	1	METE_SCHPO
22	54	19.8	936	1	FHL1_YEAST
23	54	19.8	1220	1	DP3A_THEAO
24	54	19.8	1547	1	TOR2_BOMO
25	53.5	19.6	158	1	YFPO_ECOLI
26	53.5	19.6	238	1	BPHF_ECOLI
27	53.5	19.6	375	1	Y785_METJA
28	53.5	19.6	714	1	PURL_MERTH
29	53	19.4	102	1	RS10_BACHD
30	53	19.4	102	1	RS10_STRCO
31	52.5	19.2	344	1	CHEB_CAUCR
32	52.5	19.2	478	1	FXD3_HUMAN
33	52.5	19.2	1390	1	N155_RAT

34	52.5	19.2	1391	1	N155_MOUSE
35	52	19.0	102	1	RS10_STPAN
36	52	19.0	102	1	RS10_STPN
37	52	19.0	264	1	TRA2_DROME
38	52	19.0	348	1	DJB5_HUMAN
39	52	19.0	715	1	PURL_MERTH
40	52	19.0	964	1	UL70_MCVS
41	51.5	18.9	900	1	EX7L_LACIA
42	51.5	18.9	900	1	SYA_AERPE
43	51.5	18.9	1391	1	N155_HUMAN
44	51	18.7	101	1	RS10_BACSU
45	51	18.7	101	1	RS10_CORGL

ALIGNMENTS

```

RESULT 1
TKR_DROME
ID TKR_DROME STANDARD: PRT: 753 AA.
AC P14083; 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Protein TKR.
GN TKR.
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila..
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8112827; Pubmed=3428600.
RA Haller J., Cole S., Broenner G., Jaecle H.;
RT "Dorsal and neural expression of a tyrosine kinase-related Drosophila
RT gene during embryonic development.";
RL Genes Dev. 1:862-867(1987).
CC -!- FUNCTION: POSSIBLE REGULATORY ROLE DURING DEVELOPMENT.
CC -!- CAUTION: WAS ORIGINALLY THOUGHT TO BE A KINASE ON THE BASIS OF
CC WEAK AND NON-SIGNIFICANT SIMILARITIES.
DR PIR: A27041; A27041.
DR FlyBase: FBgn003715; TKR.
DR Pfam: PF05225; HTH_Psq: 1.
FT DOMAIN 143 151 POLY-ASP.
FT DOMAIN 153 157 POLY-GLU.
FT DOMAIN 174 183 POLY-ALA.
FT DOMAIN 221 224 POLY-ASN.
FT DOMAIN 297 306 POLY-ALA.
FT DOMAIN 325 332 POLY-ALA.
FT DOMAIN 709 712 POLY-ALA.
SQ SEQUENCE 753 AA; 81021 MW; F98D3272A7DDBESE CRC64;

Query Match 25.5%: Score 69.5; DB 1; Length 753;
Best local Similarity 32.1%; Pred. No. 0.31;
Matches 17; Conservative 12; Mismatches 17; Indels 7; Gaps 1;

QY 1 KQPRKK-----RGYRQYDHEIMEEAIAMVSGKMSVKAQIGVPHSTL 46
Db 496 KMRPRKQGHAPRGPPRSWTNLTETALQHWNNKKWTSQASRIEIPYNSL 548

RESULT 2
BAB1_DROME
ID BAB1_DROME STANDARD: PRT: 977 AA.
AC Q9W0K7; 023968; O8MR78; Q9U1H7;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bric-a-brac protein 1.
GN BAB1 OR BAB OR CG9097/CG13910.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_Taxid=7227;
 RP [1]
 RP SEQUENCE FROM N.A. (ISOFORM B), FUNCTION, SUBCELLULAR LOCATION, AND
 RP TISSUE SPECIFICITY.
 RC TISSUE-Larva; and Ovary;
 RX MEDLINE=21969340; PubMed=11973274;
 RA Condorc J.L.G., Godt D., Zollman S., Chen J., Li M., Tjong S.,
 RA Cramton S.E., Sabut-Barnola I., Laski F.A.,
 RT "The bric a brac locus consists of two paralogous genes encoding
 RT BTB/POZ domain proteins and acts as a homeotic and morphogenetic
 RT regulator of imaginal development in Drosophila."
 RL Development 129:2419-2433(2002).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Champs M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Berens P.V., Berman B.P., Bhandari D., Boltslavsky S.,
 RA Botkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cline S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dubin K.J., Evangelista C.C., Ferraz C., Ferrelia S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani B., Kalush E., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekjavić G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.M., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RP [3]
 RP REVISIONS.
 RC STRAIN-Berkeley;
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review."
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RP [4]
 RP SEQUENCE FROM N.A. (ISOFORM A).
 RC STRAIN-Berkeley; TISSUE-Testis;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Garin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.,
 RT "A Drosophila full-length cDNA resource."
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 RP [5]
 RP SEQUENCE OF 99-225 FROM N.A.
 RX MEDLINE=95280944; PubMed=7760839;
 RA Chen W., Zollman S., Condorc J.L., Laski F.A.,
 RT "The bric a brac domain of bric a brac mediates dimerization in vitro."
 RL Mol. Cell. Biol. 15:3424-3429(1995).
 CC -1- FUNCTION: Probably acts as a transcriptional regulator. Required
 CC for the specification of the tarsal segment. Also involved in
 CC antenna development.
 CC -1- SUBUNIT: May form dimers.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=A;
 CC IsoId=Q9W0K7-1; Sequence=Displayed;
 CC Name=B;
 CC IsoId=Q9W0K7-2; Sequence=VSP_007015, VSP_007016;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Leg imaginal disk at the central region of the
 CC tarsus and in eye antenna disk at the basal cylinder.
 CC -1- MISCELLANEOUS: 'bric-a-brac' means 'jumble' in French (referring
 CC to the mutant ovary phenotype). DNA-binding repeat.
 CC -1- SIMILARITY: Contains 1 BTB/POZ domain.
 CC -1- SIMILARITY: Contains 1 helix-turn-helix Psq-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ252082; CAB64385.1;
 CC EMBL: AE003470; AAF47439.2;
 CC EMBL: AY122075; AAM52587.1;
 CC EMBL: U01333; AAB87052.1;
 CC Flybase: FBgn0004870; bab1.
 CC Interpro: IPR000210; BTB_POZ.
 CC Interpro: IPR007889; HTH_Psq.
 CC Pfam: PF00651; BTB; 1.
 CC Pfam: PF05225; HTH_Psq; 1.
 CC SMART: SM00225; BTB; 1.
 CC PROSITE: PS50097; BTB; 1.
 CC Nuclear protein; DNA-binding; Transcription regulation;
 CC Alternative splicing.
 CC DOMAIN 127 192
 CC FT BTB.
 CC FT H-T-H MOTIF PSQ-TYPE.
 CC FT DNA_BIND 569 614
 CC FT A.T HOOK.
 CC FT VARSPPLIC 513 526
 CC AKMNSHAMWGAT -> VSSCGDFPLANVST (in
 CC isoform A)
 CC /FTId=VSP_007015.
 CC Missing (in isoform A).
 CC /FTId=VSP_007016.
 CC K -> R (IN REF. 1).
 CC IN -> VS (IN REF. 1).
 CC A -> G (IN REF. 1 AND 5).
 CC KL -> NV (IN REF. 1).
 CC QO -> HE (IN REF. 1).
 CC D -> E (IN REF. 1).
 CC R -> M (IN REF. 1).
 CC N -> S (IN REF. 1).
 CC Q -> QOO (IN REF. 1).
 CC A -> V (IN REF. 1).
 CC FT CONFLICT 846 846
 CC FT CONFLICT 862 862
 CC FT SEQUENCE 977 AA; 103324 MW; A7676B72A43126C5 CRC64;
 SO

RC STRAIN-Berkeley;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Beres P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Deyanov A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Yao Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [3]
 RN REVISIONS.
 RP STRAIN-Berkeley;
 RC MEDLINE-22426069; PubMed-12537572;
 RX STRAIN-Berkeley;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.B., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [4]
 RN SEQUENCE OF 196-310 FROM N.A.
 RP MEDLINE-95024186; PubMed-7938017;
 RX Zolman S., Godt D., Prive G.-G., Couderc J.L., Laski F.A.;
 RA "The BTB domain, found primarily in zinc finger proteins, defines an
 RT evolutionarily conserved family that includes several developmentally
 RT regulated genes in *Drosophila*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10717-10721(1994).
 CC -1- FUNCTION: Probably acts as a transcriptional regulator. Required
 CC for the specification of the tarsal segment. Also involved in
 CC antenna development.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Leg imaginal disk at the central region of the
 CC tarsus and in eye antenna disk at the basal cylinder.
 CC -1- MISCELLANEOUS: 'birc-a-brac' means 'jumble' in French (referring to
 CC the mutant ovary phenotype).
 CC -1- SIMILARITY: Contains 1 A.T hook DNA-binding repeat.
 CC -1- SIMILARITY: Contains 1 BTB/POZ domain.
 CC -1- SIMILARITY: Contains 1 helix-turn-helix Pqg-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ252173; CAB64386.1; -
 CC EMBL: AE003470; AAF7442.2; -
 CC EMBL: U14399; AAB50834.1; -
 CC FlyBase: FBgn0025525; bab2.
 CC InterPro: IPR000637; AT hook.
 CC InterPro: IPR000210; BTB_POZ.
 CC InterPro: IPR007889; HTB_Psq.
 CC Pfam: PF00651; BTB; 1.
 CC Pfam: PF05225; HTB_Psq; 1.
 CC SMART: SM00225; BTB; 1.
 CC PROSITE: PS50097; BTB; 1.
 CC Nucleic acid binding: DNA-binding; Transcription regulation.
 CC FT DOMAIN 223 288 BTB.
 CC FT DNA_BIND 645 690 H-T-H MOTIF PSQ-TYPE.
 CC FT DNA_BIND 697 708 A-T HOOK.
 CC FT CONFLICT 858 858 A -> R (IN REF. 1).
 CC SO SEQUENCE 1067 AA; 114661 MW; 7DBFC7681D507FC0 CRC64;
 CC -----
 CC Query Match 21.6%; Score 59; DB 1; Length 1067;
 CC Best Local Similarity 31.7%; Pred. No. 12;
 CC Matches 13; Conservative 11; Mismatches 17; Indels 0; Gaps 0;
 CC -----
 CC Oy 5 KKRGRYQYDHEIMEEAIAMWSGKMSVSKAGQYGVPHST 45
 CC Db 636 RERGPLKSMRPEAAEAIFSVLKESLSQAAKRFIDIPRT 676
 CC -----
 CC RESULT 6
 CC RECD MYXXA STANDARD; PRT; 342 AA.
 CC ID RECD MYXXA STANDARD; PRT; 342 AA.
 CC AC P48291;
 CC DT 01-FEB-1996 (Rel. 33, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE RecA protein 1 (Recombinase A 1).
 CC GN RECA1.
 CC OS Myxococcus xanthus.
 CC CC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 CC CC Cystobacteriales; Myxococcaceae; Myxococcus.
 CC OX NCBI_TaxID=34;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE-95332233; PubMed-7608099;
 RA Norioka N., Hsu M.-Y., Inouye S., Inouye M.;
 RT "Two recA genes in *Myxococcus xanthus*.";
 RL J. Bacteriol. 177:4179-4182(1995).
 CC -1- FUNCTION: CAN CATALYZE THE HYDROLYSIS OF ATP IN THE PRESENCE OF
 CC SINGLE-STRANDED DNA. THE ATP-DEPENDENT UPTAKE OF SINGLE-STRANDED
 CC DNA BY DUPLEX DNA. AND THE ATP-DEPENDENT HYBRIDIZATION OF
 CC HOMOLOGOUS SINGLE-STRANDED DNAs. IT INTERACTS WITH LEXA CAUSING
 CC ITS ACTIVATION AND LEADING TO ITS AUTOCATALYTIC CLEAVAGE.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR RECA IN M.XANTHUS; RECA1
 CC SEEMS LESS FUNCTIONAL THAN RECA2.
 CC -1- SIMILARITY: Belongs to the recA family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: LA0367: AAC37000.1: -
 DR PIR: A57364: A57364.
 DR HSSP: P26345: 1G19.
 DR HAMAP: MF_00268: -; 1
 DR InterPro: IPR003593: AAA_ATPase.
 DR InterPro: IPR001553: RECA.
 DR Pfam: PF00154: RECA. 1.
 DR PRINTS: PR00142: RECA.
 DR ProDom: PD000229: RECA. 1.
 DR SMART: SM00382: AAA. 1.
 DR PROSITE: PS00321: RECA_1; 1.
 DR PROSITE: PS50162: RECA_2; 1.
 DR PROSITE: PS50163: RECA_3; 1.
 KM DNA damage: DNA recombination: SOS response: ATP-binding; DNA-binding.
 FT NF BIND 68 75 ATP (BY SIMILARITY)
 SQ SEQUENCE 342 AA; 36055 MW; 7F868E4F9FB047E4 CRC64;

Query Match 21.2%; Score 58; DB 1; Length 342;
 Best Local Similarity 33.3%; Pred. No. 4.1;
 Matches 12; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 7 RGRROYDHEIMEAIVMWSGKMSVSKAOGITGVP 42
 II : : : : : : : : : : : : : : : : : :
 DB 24 RGSVMTLGGEAREOKVAVIPSGVGVDRALGVGGYP 59

RESULT 7
 CITSY-SYN3 STANDARD; PRT; 397 AA.
 AC Q59977;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Citrate synthase (EC 2.3.3.1).
 GN GLTA OR SLT0401.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 ON NCBI_Taxid=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugita M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. I Sequence features in the 1 Mb
 RT region from map positions 648 to 928 of the genome.";
 RL DNA Res. 2:153-166(1995).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate -> citrate +
 CC COA.
 CC -1- PATHWAY: Tricarboxylic acid cycle.
 CC -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS
 CC CAPABLE OF OXIDATIVE METABOLISM.
 CC -1- SIMILARITY: Belongs to the citrate synthase family.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D64001: BAA10262.1: -
 DR PIR: S74344: S74344.
 DR HSSP: Q35354: 1A08.
 DR InterPro: IPR002020: Citrate_synth.
 DR Pfam: PF00285: citrate_synth. 1.
 DR PRINTS: PR00143: CITRISNTHASE.
 DR PROSITE: PS00480: CITRATE_SYNTHASE. 1.
 KM Tricarboxylate, tricarboxylic acid cycle; Complete proteome.
 FT ACT_SITE 266 266 BY SIMILARITY.
 FT ACT_SITE 320 320 BY SIMILARITY.
 SQ SEQUENCE 397 AA; 44830 MW; CD0A4A41F617F27E CRC64;

Query Match 21.2%; Score 58; DB 1; Length 397;
 Best Local Similarity 35.0%; Pred. No. 4.9;
 Matches 14; Conservative 7; Mismatches 15; Indels 4; Gaps 1;

QY 10 YROYDHEIMEAIVMWSGKMSVSKAOGITGVPHSTLEK 49
 I : : : : : : : : : : : : : : : : : :
 DB 3 YMMTDNEVFKEGLAGVPAKSRVSHVDGTGDI----LEVR 38

RESULT 8
 RS10_THEME STANDARD; PRT; 102 AA.
 ID RS10_THEME
 AC P38518;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S10.
 GN RPSJ OR TM1501.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 ON NCBI_Taxid=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=95095941; PubMed=8002596;
 RA Saangelantoni A.M., Bocchetta M., Cammarano P., Tiboni O.;
 RT "Phylogenetic depth of S10 and spc operators: cloning and sequencing of
 RT a ribosomal protein gene cluster from the extremely thermophilic
 RT bacterium Thermotoga maritima.";
 RL J. Bacteriol. 176:7703-7710(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima.";
 RL Nature 399:323-329(1999).
 CC -1- FUNCTION: Involved in the binding of tRNA to the ribosomes (By
 CC similarity).
 CC
 CC -1- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z21677: CAA7976.1: -
 DR EMBL: AE001798: AAD36567.1: -
 DR PIR: S40187: S40187.
 DR TIGR: TM1501: -
 DR HAMAP: MF_00508: -; 1.
 DR InterPro: IPR001848: Ribosomal_S10.
 DR InterPro: IPR005731: S10_bact.
 DR Pfam: PF00338: Ribosomal_S10. 1.
 DR PRINTS: PR00971: RIBOSOMALS10.
 DR ProDom: PD001272: Ribosomal_S10. 1.
 DR TIGRFAMs: TIGR01049: rpsJ_bact. 1.
 DR PROSITE: PS00361: RIBOSOMAL_S10. 1.
 KM Ribosomal protein; Complete proteome.
 SQ SEQUENCE 102 AA; 11637 MW; 4CB4D266F7D7C7F7 CRC64;

Query Match 20.9%; Score 57; DB 1; Length 102;
 Best Local Similarity 26.8%; Pred. No. 1.4;

Matches 15; Conservative 14; Mismatches 17; Indels 10; Gaps 2;
 QY 4 RKKRGROYDHEIMEAIAVMWS-GKMSVSKAOG-----IYGVPHSTLEK 49
 ID RS10_LACIA STANDARD; PRT; 364 AA.
 AC 026352;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan-tRNA ligase)
 GN (TPRS).
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Delta H;
 RX MEDLINE=96037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Albrecht T., Bascichirad R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McQuigall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mo J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics.";
 CC J. Bacteriol. 179:7135-7155(1997).
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -I- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE000812; AAB84757.1;
 DR PIR: E69131; E69131.
 DR HAMAP: MF_00140; -; 1.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; Trp_tRNA-synt_1b.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASTYTHRP.
 DR TIGRFAMs: TIGR00233; tprs; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 FT SITE 70 "HIGH" REGION.
 FT SITE 251 "MSKS" REGION.
 SO SEQUENCE 364 AA; 41301 MW; C2F348903338F61D CRC64;

Query Match 20.7%; Score 56.5; DB 1; Length 364;
 Best Local Similarity 36.1%; Pred. No. 7.1;
 Matches 13; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

QY 10 RYRDHEIMEAIAVMWSGKMSVSKAOGIYGVPHST 45
 ID RS10_LACIA STANDARD; PRT; 364 AA.
 AC 026352;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan-tRNA ligase)
 GN (TPRS).
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Delta H;
 RX MEDLINE=96037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Albrecht T., Bascichirad R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McQuigall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mo J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics.";
 CC J. Bacteriol. 179:7135-7155(1997).
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -I- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE000812; AAB84757.1;
 DR PIR: E69131; E69131.
 DR HAMAP: MF_00140; -; 1.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; Trp_tRNA-synt_1b.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASTYTHRP.
 DR TIGRFAMs: TIGR00233; tprs; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 FT SITE 70 "HIGH" REGION.
 FT SITE 251 "MSKS" REGION.
 SO SEQUENCE 364 AA; 41301 MW; C2F348903338F61D CRC64;

RESULT 10
 RS10_LACIA
 ID RS10_LACIA STANDARD; PRT; 102 AA.
 AC 09CDM1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S10.
 GN RPSJ OR IL2100.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 lactis ssp. lactis IL1403.";
 RL Genome Res. 11:731-753(2001).
 CC -I- FUNCTION: Involved in the binding of tRNA to the ribosomes (By
 CC similarity).
 CC -I- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE006438; AAK06198.1;
 DR PIR: D86887; D86887.
 DR HAMAP: MF_00508; -; 1.
 DR InterPro: IPR001848; Ribosomal_S10.
 DR InterPro: IPR005731; S10_bact.
 DR Pfam: PF00338; Ribosomal_S10; 1.
 DR PRINTS: PR00971; RIBOSOMALS10.
 DR Prodom: PD001272; RIBOSOMALS10.
 DR TIGRFAMs: TIGR01049; rpsJ_bact; 1.
 DR PROSITE: PS00361; RIBOSOMAL_S10; 1.
 KW Ribosomal protein; Complete proteome.
 SO SEQUENCE 102 AA; 11742 MW; 15D8A683A2B682A0 CRC64;

Query Match 20.5%; Score 56; DB 1; Length 102;
 Best Local Similarity 25.0%; Pred. No. 1.9;
 Matches 14; Conservative 13; Mismatches 19; Indels 10; Gaps 1;

QY 4 RKKRGROYDHEIMEAIAVMWSGKMSVSKAOGIYGVPHSTLEK 49
 ID RS10_LACIA STANDARD; PRT; 102 AA.
 AC 0927K6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S10.
 GN RPSJ OR LMO2633 OR LIN2782.
 OS Listeria monocytogenes, and
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639; 1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=L.monocytogenes, and L.innocua;
 RX STRAIN=EGD-e / Serovar 1/2a, and CLIP 11262 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;

RESULT 11
 RS10_LISMO
 ID RS10_LISMO STANDARD; PRT; 102 AA.
 AC 0927K6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S10.
 GN RPSJ OR LMO2633 OR LIN2782.
 OS Listeria monocytogenes, and
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639; 1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=L.monocytogenes, and L.innocua;
 RX STRAIN=EGD-e / Serovar 1/2a, and CLIP 11262 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;

```

RA glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charlot A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA Doman E., Dominguez-Bernal G., Duchaud E., Durat L., Dussuret O.,
RA Ertan K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gauthier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kunaprat G.,
RA Madueno E., Maitounan A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.,
RA Science 294:849-852(2001).
CC -1- FUNCTION: Involved in the binding of tRNA to the ribosomes (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL596173; CAC98008.1; -
CC DR EMBL: AL591983; CAD00711.1; -
CC DR PIR: AH1779; AH1779.
CC DR PIR: AH1403; AH1403.
CC DR Listlist: LM002633; -
CC DR Listlist: LINO2782; -
CC DR HAMAP: MF_00508; -; 1.
CC DR InterPro: IPR001848; Ribosomal_S10.
CC DR InterPro: IPR005731; S10_bact.
CC DR Pfam: PF00338; Ribosomal_S10.1.
CC DR PRINTS: PR00971; RIBOSOMAL_S10.
CC DR PRODOM: PD001272; Ribosomal_S10.1.
CC DR TIGRFAMs: TIGR01049; rpsJ_bact; 1.
CC DR PROSITE: PS00361; RIBOSOMAL_S10.1.
CC KW Ribosomal protein; Complete proteome.
CC SO SEQUENCE 102 AA; 11682 MW; 7E4B186EA533CCDB CRC64;

Query Match 20.5%; Score 56; DB 1; Length 102;
Best Local Similarity 25.0%; Pred. No. 1.9;
Matches 14; Conservative 13; Mismatches 19; Indels 10; Gaps 1;

OY 4 KKKRGROYDHEIMEAIVM-----MSGKMSVSKAOGIYGVPHSTLEYK 49
DB 4 OKIRIRLKAYDHRILDSASEKIVETAKRSGASVSGPIPLPTKRSIYVLRVHKYK 59

RESULT 12
RS10_HELPY STANDARD; PRT; 104 AA.
AC P56017;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S10.
GN rpsJ OR HP1320 OR JHP1240.
OS Helicobacter pylori (Campylobacter pylori), and
OS Helicobacter pylori 399 (Campylobacter pylori 399).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
NCBI_TaxID=210; 85963;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,

```

```

RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watney L., Mallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RT Nature 388:539-547(1997).
RA SEQUENCE FROM N.A.
RP STRAIN=J99;
RC MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carnel G.,
RA Tummino P.J., Carnuso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RT Nature 397:176-180(1999).
CC -1- FUNCTION: Involved in the binding of tRNA to the ribosomes (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF000633; AAD08359.1; -
CC DR EMBL: AE001547; AAD06791.1; -
CC DR PIR: H64684; H64684.
CC DR TIGR: HP1320; -
CC DR HAMAP: MF_00508; -; 1.
CC DR InterPro: IPR001848; Ribosomal_S10.
CC DR InterPro: IPR005731; S10_bact.
CC DR Pfam: PF00338; Ribosomal_S10.1.
CC DR PRINTS: PR00971; RIBOSOMAL_S10.
CC DR PRODOM: PD001272; Ribosomal_S10.1.
CC DR TIGRFAMs: TIGR01049; rpsJ_bact; 1.
CC DR PROSITE: PS00361; RIBOSOMAL_S10.1.
CC KW Ribosomal protein; Complete proteome.
CC SO SEQUENCE 104 AA; 11918 MW; 7EEA60A965B41CE1 CRC64;

Query Match 20.3%; Score 55.5; DB 1; Length 104;
Best Local Similarity 29.8%; Pred. No. 2.2;
Matches 14; Conservative 12; Mismatches 20; Indels 1; Gaps 1;

OY 5 KKKRGROYDHEIMEA-IVMWSGKMSVSKAOGIYGVPHSTLEYK 50
DB 3 KIRLKLKAYDHRVLDKRSVAIVAEVAKRSGSEIRGPILPTKRNKRYTV 49

RESULT 13
Y769_METUA STANDARD; PRT; 169 AA.
AC O58179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0769.
GN M0769.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
NCBI_TaxID=2190;
RX MEDLINE=96337999; PubMed=8688087;

```

DR	PIR: C71517: C71517.	AAA,ATPase.
DR	InterPro: IPR003593; ABC_Transporter.	
DR	InterPro: IPR003439; ABC_Transporter.	
DR	PIfam: PF00005; ABC_tran; 1.	
DR	SMART: SM00382; AAA; 1.	
DR	PROSITE: PS00211; ABC_TRANSPORTER_1; 1.	
DR	PROSITE: PS50893; ABC_TRANSPORTER_2; 1.	
KW	Hypothetical protein; Transport; Inner membrane; ATP-binding;	
FT	Complete proteome.	
FT	NP_BIND	39
SO	SEQUENCE	236 AA; 26193 MW; DE74774507950365 CRC64;
QY	8 GRYRQDHEIMEEAIANV 25	
Db	109 GRVTOEDHEAEAEALTLTV 126	
RESULT 15		
STH_STAMM	STANDARD;	PRT; 420 AA.
ID	SYN_STAMM	
AC	O32422;	
DT	15-DEC-1998 (Rel. 37, Created)	
DT	15-DEC-1998 (Rel. 37, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	HistHyl-trNA synthetase (EC 6.1.1.21) (Histidine--trNA ligase) (HISRS)	
GN	HISS OR SAV1631 OR SA1457 OR MW1581.	
OS	Staphylococcus aureus (strain M50 / ATCC 700699),	
OS	staphylococcus aureus (strain N315),	
OS	Staphylococcus aureus (strain MM2), and	
OS	Staphylococcus aureus.	
CC	Bacteria; Firmicutes; Bacillales; Staphylococcus.	
OX	NCBI_TaxID=158878, 158879, 196620, 1280;	
RN	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=M50 / ATCC 700699, and N315;	
RX	MEDLINE=21311952; PubMed=11418146;	
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,	
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,	
RA	Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,	
RA	Mitubuti-Oi Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,	
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,	
RA	Kanehisa M., Yanaoishi A., Oshima K., Furuya K., Yoshino C., Shiba T.,	
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;	
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus	
RT	aureus."	
RL	Lancet 357:1225-1240(2001).	
RN	[2]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=MM2;	
RX	MEDLINE=22040718; PubMed=12044378;	
RA	Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,	
RA	Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,	
RA	Yamamoto K., Hiramatsu K.;	
RT	"Genome and virulence determinants of high virulence community-	
RT	acquired MRSA."	
RL	Lancet 359:1819-1827(2002).	
RN	[3]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=SR17238;	
RX	MEDLINE=97474251; PubMed=9335275;	
RA	Fujimura T., Murakami K.;	
RT	"Increase of methicillin resistance in Staphylococcus aureus caused	
RT	by deletion of a gene whose product is homologous to lytic enzymes."	
RL	J Bacteriol. 179:6294-6301(1997).	
RN	[4]	
RP	X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).	
RX	MEDLINE=99425117; PubMed=10493797;	
RA	Qiu X., Janson C.A., Blackburn M.N., Chohan I.K., Hibbs M.,	

THIS PAGE BLANK (USPTO)


```

A>Title: The Drosophila pipsqueak gene encodes a nuclear BTB-domain-containing protein
A.Reference number: Z22972; MUID:96632302; PMID:8674425
A.Accession: T45461
A>Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: mRNA
A.Residues: 1-355, 'E', 357-1005, 'H', 1007-1020, 'Q', 1021-1061, 'ERS' <HOR>
A.Cross-References: EMBL:U048358; NID:g1203906; PIDN:AAC47153.1; PTD:g1203907
C.Experimental source: tissue type ovarian
C.Genetics:
A.Gene: pipsqueak; psq
A.Map position: II
A.Introns: 427/3
C.Function:
A.Description: required for establishing polarity of the developing egg chamber
C.Superfamily: BRCre-2 protein; POZ domain homology
F:21-123/Domain: POZ domain homology <POZ>

Query Match          32.4%; Score 88.5; DB 2; Length 1085;
Best Local Similarity 34.7%; Pred. No. 0.0062;
Matches   17; Conservative    14; Mismatches   17; Indels      1; Gaps     1;

Oy      5 KKRGRDYDHIMEAIAMVMSGKMSKSGOGIYGVPHSTLEKYKKR 53
       |||::: : ::||:::||||:::||::||| |::|
Db      773 KKGGTKWNEDALONALELRSGDISANKSKAFGPSLT-LYKIARR 820

RESULT 3
A27041
tyrosine kinase-related protein - fruit fly (Drosophila melanogaster)
C.Species: Drosophila melanogaster
C.Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 04-Feb-2000
C.Accession: A27041
R.Haller, J.; Cole, S.; Broenner, G.; Jaeckle, H.
Genes Dev. 1, 862-867, 1987
A.Title: Dorsal and neural expression of a tyrosine kinase-related Drosophila gene during
A.Reference number: A27041; MUID:88112827; PMID:3428600
A.Accession: A27041
A.Status: not compared with conceptual translation
Molecule type: DNA
A.Residues: 1-753 <HAL>
C.Genetics:
A.Gene: dtkr
A.Cross-references: FlyBase:FBgn0003715
A.Map position: 2R,60P1
A.Introns: 453/1; 497/1
C.Keywords: autophosphorylation; glycoprotein; phosphoprotein
F:9,65,187,223,224,250,611,660/Binding site: carbohydrtse (Asn) (covalent) #status predicted
F:744/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match          25.5%; Score 69.5; DB 2; Length 753;
Best Local Similarity 32.1%; Pred. No. 1.1;
Matches   17; Conservative    12; Mismatches   17; Indels      7; Gaps     1;

Oy      1 KORPK-----RGRRYQYDHIEAEIAIMVMSGKSNSKAOCIGYPHSTL 46
       ||| | | | | | | | | | | | | | | | | | | | | | |
Db      496 KMPRKCGHSAPRCGRPRSRMTNELTEALQHVMWKKMTTSQSASRLFGIPNSTL 548

RESULT 4
G65007
hypothetical protein b2345 - Escherichia coli (strain K-12)
C.Species: Escherichia coli
C.Date: 12-Sep-1997 #sequence_revlsion 17-Sep-1997 #text_change 01-Mar-2002
C.Accession: G65007
R.Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cai
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A.Title: The complete genome sequence of Escherichia coli K-12.
A.Reference number: A64720; MUID:97426617; PMID:9278503
A.Accession: G65007
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
Molecule type: DNA
A.Residues: 1-352 <BLAT>
```

```

A:Cross-References: GB:AEO00323; GB:U00096; NID:g1788664; PIDN:AAC75405.1; PID:g1788664
A:Experimental source: strain K-12, substrain M6165
C:Superfamily: Escherichia coli hypothetical protein b2345

Query Match      22.3%; Score 61; DB 2; Length 352;
Best Local Similarity 28.8%; Pred. No. 5.8;
Matches   17; Conservative    15; Mismatches   15; Indels   12; Gaps    3;

QY          5 KKRGVROYDHEI-----MEEAIAMWSGKMSYSKAQGIGVPH-STLEYKKER 53
Db         271 KRGRAIRYPDVDADAKKYCDHLKEIVSTLVGNKINMGHSOKIY-KHFKDLSEKIEER 327

RESULT 5
BA0628
resolvase homolog - Enterococcus faecium transposon Tn1546
C:Species: Enterococcus faecium
C>Date: 21-Sep-1993 #sequence_revision 07-Oct-1994 #text_change 21-Aug-1998
C:Accession: BA0628
R:Arthur, M.; Molinas, C.; Depardieu, F.; Courvalin, P.
J. Bacteriol. 175, 117-127, 1993
A>Title: Characterization of Tn1546, a tn-related transposon conferring glycopeptide
A:Reference number: A40628; MUID:93106944; PMID:8380148
A:Contents: BM4147, pP816
A:Accession: BA0628
A:Molecule type: DNA
A:Residues: 1-191 <ART>
A:Cross-References: GB:M97297
A>Note: sequence extracted from NCBI backbone (NCBIN:121238, NCBIP:121242)
C:Superfamily: transposase repressor
C:Keywords: DNA binding

Query Match      21.6%; Score 59; DB 2; Length 191;
Best Local Similarity 29.6%; Pred. No. 5.3;
Matches   16; Conservative    13; Mismatches   23; Indels   2; Gaps    1;

QY          1 KOPRRKRORYOY--DHEIMEATIAMWSGKSMSYSKAQGIYGVPHTLEKYKE 52
Db        135 KREGFRGLRKHKHNAGMNYAVKLTYEGMTVNQCETITVSASLYRRLSE 188

RESULT 6
S51715
hypothetical protein - Chlamydia trachomatis (fragment)
C:Species: Chlamydia trachomatis
C>Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 11-Jan-2000
C:Accession: S51715
R:Pohlner, J.
submitted to the EMBL Data Library, April 1994
A:Description: Sequence of the recA gene from Chlamydia trachomatis L2.
A:Reference number: S51714
A:Accession: S51715
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <POH>
A:Cross-References: EMBL:Z32530; NID:g607060; PIDN:CAA83540.1; PID:g607062
C:Superfamily: Chlamydia hypothetical protein CT648

Query Match      21.6%; Score 59; DB 2; Length 419;
Best Local Similarity 32.7%; Pred. No. 13;
Matches   17; Conservative    10; Mismatches   13; Indels   12; Gaps    3;

QY          1 KOPRRKGRRORYDHEIMEEAIAWMSGK-----MSVSRAQGIYGVP 43
Db        227 KGGLRSRGKV-YD-PLTKEEICCVNVGGKKOVIYGYALIEROIHYGVPH 275

RESULT 7
F71488
hypothetical protein CT648 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 10-Dec-1999
C:Accession: F71488

```

```

Science 282, 754-759, 1998
A>Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: F71488
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <ARN>
A:Cross-references: GB:AE001335; GB:AE001273; NID:g3329091; PIDN:AAC68825.1; PID:g332909
C:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT648
C:Superfamily: Chlamydia hypothetical protein CTr648

Query Match .21.6% Score 59; DB 2; Length 424;
Best Local Similarity 32.7%; Pred. No. 13;
Matches 17; Conservative 10; Mismatches 13; Indels 12; Gaps 3;

OY 1 KQPKKRGRYQDYHEIMEEALAMVMSGK-----MVSXKQGCTGYCPH 43
||:::||:| | : || | ::|||
Db 232 KGKLRSGLKY--YD-PLTKKEIALCVNGSKROYIYGRYAETROIVHGVPH 280

RESULT 8
A57364
recombination protein recA - Myxococcus xanthus
N:Alternate names: recombinae A
C:Species: Myxococcus xanthus
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 02-Feb-2001
C:Accession: A57364
R:Norioka, N.; Hsu, M.Y.; Inouye, S.; Inouye, M.
J. Bacteriol. 177, 4179-4182, 1995
A>Title: Two recA genes in Myxococcus xanthus.
A:Reference number: A57364; MUID:95332233; PMID:7608099
A:Accession: A57364
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <RES>
A:Cross-references: GB:I40367; NID:g848993; PIDN:AAC37000.1; PID:g848994
C:Genetics:
A:Gene: recA
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:68-75/Region: nucleotide-binding motif A (P-loop)
F:142-147/Region: nucleotide-binding motif B
F:74/Binding site: ATP (Lys) #status predicted

Query Match 21.2% Score 58; DB 2; Length 342;
Best Local Similarity 33.3%; Pred. No. 14;
Matches 12; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

OY 7 RGRYRDYEIMEEALAMVMGSKMSKAOGTGV 42
|| ||: |: | : | : | : | : | : |
Db 24 RGSYMTLGGEARERKVAVIPSGSYCDRALGVGP 59

RESULT 9
S74344
citrate synthase gltA - Synecchocystis sp. (strain PCC 6803)
N:Alternate names: hypothetical protein sl10401
C:Species: Synecchocystis sp.
A:Variety: PCC 6803
C:date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
C. K., Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Matsubae, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-116, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74344
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
```

[illegible]

RESULT 12

H71033
hypothetical protein PH1560 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 28-Jul-2000
C:Accession: H71033
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekitani, S.; Ohnuki, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:96344137; PMID:9679194
A:Accession: H71033
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-140 <KAM>
A:Cross-references: GB:AF000006; NID:93236133; PIDN:BA30672.1; PID:93257989
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1560
C:Superfamily: Pyrococcus abyssi hypothetical protein PAB2413

Query Match 21.1%; Score 57.5; DB 2; Length 140;
Best Local Similarity 35.4%; Pred. No. 5.8;
Matches 17; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

OY 5 KKRGRYQYDHEIMEAIAWMSK--MSVSKAGIYGVPHSTLEKY 50
| | | | | : | : | | | : | | | : | | | : |
Db 85 KGRGRPKYSTRTLEVINEMLRSGKTPREISK---LGIPIRTLYYL 129

RESULT 13

S40187
ribosomal protein S10 - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 21-Jul-2000
C:Accession: S40187; E72250
R:Sanangelantoni, A.; Tiboni, O.
Submitted to the EMBL Data Library, February 1993
A:Reference number: S37489
A:Accession: S40187
A:Molecule type: DNA
A:Residues: 1-102 <SAN>
A:Cross-references: EMBL:Z21677; NID:9437921; PIDN:CA47976.1; PID:9437922
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: E72250
A:Molecule type: DNA
A:Residues: 1-102 <ARN>
A:Cross-references: GB:AE001798; GB:AE000512; NID:94982033; PIDN:AD36567.1; PID:9498206
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: rpsJ
C:Superfamily: Escherichia coli ribosomal protein S10
C:Keywords: protein biosynthesis; ribosome

Query Match 20.9%; Score 57; DB 2; Length 102;
Best Local Similarity 26.8%; Pred. No. 4.7;
Matches 15; Conservative 14; Mismatches 17; Indels 10; Gaps 2;

OY 4 RKRGRYQYDHEIMEAIAWMS--GKMSVSKAG-----TYGVPHSTLEKY 49
| | | | | : | | | | | : | | | : | | | : |
Db 4 OKIRIKLAKYDHEILDSAKKIYEVAKSTNKGSPILPRTLYCYLRSFPMKH 59

RESULT 14

E69131

tryptophan-tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain De N:Alternate names: tryptophanyl-tRNA synthetase
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 03-Jun-2002
C:Accession: E69131
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicalire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivanli, K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:96037514; PMID:93714463
A:Accession: E69131
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-364 <MTH>
A:Cross-references: GB:AF000812; GB:AE000666; NID:92621298; PIDN:AB84757.1; PID:9262
A:Experimental source: strain Delta H

Query Match 20.7%; Score 56.5; DB 2; Length 364;
Best Local Similarity 36.1%; Pred. No. 23;
Matches 13; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

OY 10 YROYDHEIMEAIAWMSGKMSVSKAGIYGVPHST 45
| | | | | : | | | | | : | | | : | | | : |
Db 142 YLOSEN-LMWEDLAVLAGKVFNEFLAIVGFTGST 176

RESULT 15

F75406
histidyl-tRNA synthetase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: F75406
R:White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Shen, M.; Vamathevan, J.T.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75406
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-427 <WHI>
A:Cross-references: GB:AE001981; GB:AE000513; NID:96459097; PIDN:AF10919.1; PID:9645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1349
C:Superfamily: histidine-tRNA ligase; histidine-tRNA ligase homology

Query Match 20.7%; Score 56.5; DB 2; Length 427;
Best Local Similarity 33.3%; Pred. No. 27;
Matches 15; Conservative 13; Mismatches 10; Indels 7; Gaps 2;

OY 4 RKRGRYQYDHEIMEAIAW---MSGKMSVSKAGIYCV 41
| | | | | : | | | | | : | | | : | | | : |
Db 125 RHQGRYQYDHEIVLGSADALVDALIAWVAVQKGLRCV 169

Search completed: August 4, 2003, 15:07:32
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 14:54:34 ; Search time 83 seconds

(without alignments)
101.356 Million cell updates/sec

Title: US-10-016-768a-2

Perfect score: 273
Sequence: 1 KOPRRKRRGRROYDHEIMEE.....KAGQIVGPHSTLEKYKER 53

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	273	100.0	53	23	AAE24592 Human E93 program
2	273	100.0	442	23	AAE24371 Human E93 program
3	273	100.0	630	22	ABG17942 Human E93 program
4	234	85.7	104	23	ABP32451 Human ORF1424 prot
5	233.5	85.5	54	23	AAE24593 Fish E93 programme
6	229	83.9	53	23	AAE24594 Mouse E93 programme
7	165	60.4	53	23	AAE24370 Fruit fly E93 prog
8	165	60.4	1140	22	ABR71145 Drosophila melanog
9	165	60.4	1165	23	AAE24372 Fruit fly E93 prog

10	163	59.7	53	23	AAE24595
11	88.5	32.4	1046	22	ABR67028
12	88.5	32.4	1064	22	ABR59068
13	70.5	25.8	661	22	ABR53113
14	69.5	25.5	1046	22	ABR59307
15	63	23.1	457	22	ABR68768
16	62	22.7	325	24	ABP56413
17	61	22.3	352	22	ACG98384
18	61	22.3	494	23	ABP26391
19	61	22.3	494	23	ABP29724
20	59.5	21.8	263	22	ABR64196
21	59	21.6	191	13	AAE24294
22	59	21.6	434	20	AAE37533
23	59	21.6	484	22	ABR68769
24	58	21.2	141	22	AAO05999
25	58	21.2	364	18	AAW34643
26	58	21.2	364	23	ABR82281
27	58	21.2	364	23	ABG70786
28	58	21.2	364	23	ABG70828
29	58	21.2	364	23	ABG70919
30	58	21.2	364	23	ABG31855
31	58	21.2	364	24	ABG74442
32	58	21.2	364	24	ABG73912
33	58	21.2	364	24	ABG73965
34	57.5	21.1	957	22	ABR64362
35	57	20.9	298	21	ABR07753
36	56.5	20.7	2027	22	AAE10987
37	56	20.5	102	23	ABR55454
38	56	20.5	102	23	ABR49361
39	56	20.5	1693	22	AAW80141
40	56	20.5	1723	22	AAW79157
41	55.5	20.3	104	22	AAU35850
42	55.5	20.3	104	22	AAU36004
43	55.5	20.1	828	22	ABR57774
44	55	20.3	204	22	AAE91636
45	55	20.1	236	20	AAV37820

ALIGNMENTS

RESULT 1	
ID	AAE24592
AAE24592	standard; Protein: 53 AA.
AC	AAE24592:
DT	04-OCT-2002 (first entry)
XX	Human E93 programmed cell death modulating protein conserved domain.
XX	Human; cancer; programmed cell death modulating protein; adenocarcinoma;
KW	cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;
KW	neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;
KW	Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;
KW	aplastic anaemia; ischemic injury; myocardial infarction; stroke;
KW	reperfusion injury; toxin-induced disease; genetic immunodeficiency;
KW	vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;
KW	myeloma; nootropic; vasotropic; immunostimulant; cerebroprotective;
XX	cardiant; E93 protein.
OS	Homo sapiens.
PN	WO200234882-A2.
PD	02-MAY-2002.
XX	29-OCT-2001; 2001WO-US48053.
PF	27-OCT-2000; 2000US-243865P.
XX	(UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.
PA	
XX	

Nematode E93 prog
Drosophila melanog
Drosophila melanog
Drosophila melanog
Drosophila melanog
Drosophila melanog
Thermus thermophil
Escherichia coli p
Streptococcus poly
Streptococcus poly
Drosophila melanog
Resolvase encoded
Chlamydia trachom
Drosophila melanog
Human polypeptide
Thermococcus alca
T. alcaliphilus AE
T. alcaliphilus AE
T. alcaliphilus AE
Thermococcus alca
Thermococcus alca
T. alcaliphilus al
Thermococcus alca
Thermococcus alca
Drosophila melanog
An oocyte maturat
Mouse presentin/C
Lactococcus lactis
Listeria monocytog
Human protein SEQ
Human protein SEQ
Helicobacter pylor
Helicobacter pylor
Drosophila melanog
C glutamicum prote
Amino acid sequenc

Pt	Baehrecke EH:
Xl	WPI: 2002-479717/51.
Dx	
Pt	Noval programmed cell death modulating proteins, useful for treating or preventing disorders associated with abnormal cell proliferation and apoptosis such as cancer, stroke, Parkinson's disease, myocardial infarction -
Pf	-
Xx	
Px	Claim 1; Fig 1; 88pp: English.
Xx	
Cc	The present invention relates to novel programmed cell death modulating proteins and polynucleotides encoding such proteins. Sequences of the invention are useful to screen potential cellular apoptosis inhibiting compounds to determine their use as therapeutic agents for treatment of diseases associated with increased programmed cell death. They are also useful for treating or preventing disorders associated with decrease in apoptosis. Programmed cell death modulating sequences are useful for treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma. Inhibition of the activity of the sequences of the invention are useful for treating disorders associated with increase in cell death or apoptosis such as acquired immunodeficiency syndrome (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis pigmentosa, Parkinson's disease and cerebellar degeneration), ischemic injuries (e.g., myocardial infarction, stroke, reperfusion injury), myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced diseases and other infectious or genetic immunodeficiencies. Sequences of the invention are used as vaccines and in gene therapy. The present sequence is human E93 programmed cell death modulating protein conserved domain.
Cc	
Cc	
Cc	
Sq	Sequence 53 AA: Query Match 100.0%; Score 273; DB 23; Length 53; Best Local Similarity 100.0%; Pred. No. 2.5e-33;
Xx	Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 KQPKKRGRYQYDHEIMEALVAVMSKMVSNAQGTYGVPHSTLEKKVER 53 1 KQPKKRGRYQYDHEIMEALVAVMSKMVSNAOGTIVGPHSTLEKKVER 53
Dd	1 KQPKKRGRYQYDHEIMEALVAVMSKMVSNAOGTIVGPHSTLEKKVER 53
Rstult 2	
Aae24371	ID AAE24371 standard; Protein: 442 AA. XX AC AAE24371; DT XX XX 04-OCT-2002 (first entry) XX Human E93 programmed cell death modulating protein.
Kw	Human: cancer; programmed cell death modulating protein; adenocarcinoma; cellular apoptosis; leukemia; acquired immunodeficiency syndrome; AIDS; neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa; Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration; aplastic anemia; ischemic injury; myocardial infarction; stroke; reperfusion injury; toxin-induced disease; genetic immunodeficiency; vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective; myeloma; nootropic; vasotropic; immunostimulant; cerebroprotective; cardiant; E93 protein.
Os	Homo sapiens.
FH	Key Location/Qualifiers FT Domain 353..405 /note="Conserved domain"
Pf	/note="Conserved domain"
Xx	
Pn	WO200234862-A2.
Xx	
Dd	02-MAY-2002.
Xx	
Pf	29-OCT-2001; 2001WO-US48053.

XX	27-OCT-2000; 2000US-243865P.
PR	(UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.
PA	Baehrecke EH;
PI	WPI; 2002-479717/51.
XX	
DR	
XX	
PT	Novel programmed cell death modulating proteins, useful for treating or
PT	preventing disorders associated with abnormal cell proliferation and
PT	apoptosis such as cancer, stroke, Parkinson's disease, myocardial
PT	infarction -
XX	
PS	Claim 1; Fig 4; 88pp; English.
XX	
XX	The present invention relates to novel programmed cell death modulating
CC	proteins and polynucleotides encoding such proteins. Sequences of the
CC	invention are useful to screen potential cellular apoptosis inhibiting
CC	compounds to determine their use as therapeutic agents for treatment of
CC	diseases associated with increased programmed cell death. They are also
CC	useful for treating or preventing disorders associated with decrease in
CC	apoptosis. Programmed cell death modulating sequences are useful for
CC	treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,
CC	melanoma, myeloma. Inhibition of the activity of the sequences of the
CC	invention are useful for treating disorders associated with increase
CC	in cell death or apoptosis such as acquired immunodeficiency syndrome
CC	(AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis
CC	pigmentosa, Parkinson's disease and cerebellar degeneration), ischemic
CC	injuries (e.g., myocardial infarction, stroke, reperfusion injury),
CC	myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced
CC	diseases and other infectious or genetic immunodeficiencies. Sequences
CC	of the invention are used as vaccines and in gene therapy. The present
CC	sequence is human E93 programmed cell death modulating protein.
XX	
SQ	Sequence 442 AA;
XX	
Query Match	100.0%; Score 273; DB 23; Length 442;
Best Local Similarity	100.0%; Pred. No. 4e-32;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 KOPPKRRRYQYDHEIMEAIAMVMSGKMSVSKAOGYGVPHSTLEKYKER 53
DB	353 KOPKKRRRYQYDHEIMEAIAMVMSGKMSVSKAOGYGVPHSTLEKYKER 405
XXXX	
RESULT 3	
ABG17942	
ID	ABG17942 standard; Protein: 630 AA.
XX	
AC	ABG17942;
XX	
DT	18-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #17933.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	food supplement; medical imaging; diagnostic; genetic disorder.
OS	Homo sapiens.
XX	
WO200175067-A2.	
PN	
PD	11-OCT-2001.
XX	
PE	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.
DR N-PSDB: AAS82129.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX
PS Claim 20, SEQ ID NO 48301, 103pp: English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX
SQ Sequence 630 AA:
Query Match 100.0%; Score 273; DB 22; Length 630;
Best Local Similarity 100.0%; Pred. No. 6,4e-32;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KOPRRKRGRRYQYDHEIMEAIVMVGSKMSVSKAGIYGVPHSTLEYKVKER 53
Db 541 KOPRRKRGRRYQYDHEIMEAIVMVGSKMSVSKAGIYGVPHSTLEYKVKER 593
RESULT 4
ID ABB32451 standard; Protein: 104 AA.
XX
AC ABB32451:
XX
DT 09-JUL-2002 (first entry)
XX
DE Human ORF1424 protein, SEQ ID NO:2848.
XX
XX Human: ORF: open reading frame; ORFX: drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antiproliferative; antidiabetic; cytostatic; neutrophic;
KW neuroprotective; antithrombotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
XX Homo sapiens.
XX
OS MO200190366-A2.
XX
PN 29-NOV-2001.
XX

XX
XX 24-MAY-2001; 2001WO-US17076.
PF
XX
XX
PR 24-MAY-2000; 2000US-206690P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Leach MD, Shinkets RA:
XX
XX WPI: 2002-106200/14.
DR N-PSDB: ABB76477.
XX
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation
XX
XX
PS Claim 10; Page 971-972; 2508pp: English.
XX
XX Sequences ABB31028-ABB35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABB75054-
CC ABB75587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes. In the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.
XX
XX
SQ Sequence 104 AA:
Query Match 85.7%; Score 234; DB 23; Length 104;
Best Local Similarity 84.9%; Pred. No. 4,4e-27;
Matches 45; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
OY 1 KOPRRKRGRRYQYDHEIMEAIVMVGSKMSVSKAGIYGVPHSTLEYKVKER 53
Db 8 KOPRRKRGRRYQYDHEIMEAIVMVGSKMSVSKAGIYGVPHSTLEYKVKER 60
RESULT 5
ID AAE24593 standard; Protein: 54 AA.
XX
XX AAE24593:
XX
AC AAE24593:
XX
DT 04-OCT-2002 (first entry)
XX

[illegible]

DE	Mouse E93 programmed cell death modulating protein conserved domain.	
XX		
XX	Mouse; cancer; programmed cell death modulating protein; adenocarcinoma;	
KW	cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;	
KW	neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;	
KW	Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;	
KW	aplastic anaemia; ischaemic injury; myocardial infarction; stroke;	
KW	reperfusion injury; toxin-induced disease; genetic immunodeficiency;	
KW	vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;	
KW	myeloma; neurotropic; vasculotropic; immunostimulant; cerebroprotective;	
KW	cardiant; E93 protein.	
OS	Mus musculus.	
XX		
PN	WO200234882-A2.	
XX		
PD	02-MAY-2002.	
XX		
PF	29-OCT-2001; 2001WO-US48053.	
XX		
PR	27-OCT-2000; 2000US-243865P.	
XX		
XX	(UYWA-) UNIV MARIAND BIOTECHNOLOGY INST.	
PA		
PI	Baehrecke EH;	
XX		
DR	WPI; 2002-479717/51.	
XX		
PT	Novel programmed cell death modulating proteins, useful for treating or	
PT	preventing disorders associated with abnormal cell proliferation and	
PT	apoptosis such as cancer, stroke, Parkinson's disease, myocardial	
PT	infarction	
XX		
PS	Claim 1; Fig 1; 88pp; English.	
XX		
CC	The present invention relates to novel programmed cell death modulating	
CC	proteins and polynucleotides encoding such proteins. Sequences of the	
CC	invention are useful to screen potential cellular apoptosis inhibiting	
CC	compounds to determine their use as therapeutic agents for treatment of	
CC	diseases associated with increased programmed cell death. They are also	
CC	useful for treating or preventing disorders associated with decrease in	
CC	apoptosis. Programmed cell death modulating sequences are useful for	
CC	treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,	
CC	melanoma, myeloma. Inhibition of the activity of the sequences of the	
CC	invention are useful for treating disorders associated with increase	
CC	in cell death or apoptosis such as acquired immunodeficiency syndrome	
CC	(AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis	
CC	pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic	
CC	injuries (e.g., myocardial infarction, stroke, reperfusion injury),	
CC	myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced	
CC	diseases and other infectious or genetic immunodeficiencies. Sequences	
CC	of the invention are used as vaccines and in gene therapy. The present	
CC	sequence is mouse E93 programmed cell death modulating protein conserved	
CC	domain.	
CC		
XX		
SQ	Sequence 53 AA;	
XX		
Query Match	83.9%; Score 229; DB 23; Length 53;	
Best Local Similarity	81.1%; Pred. No. 1e-26;	
Matches	43; Conservative 6; Mismatches 4; Indels 0; Gaps 0;	
Db	1 KOPRRKRRYQYDHEIMEEAIAMWSGSKMSVSKAAGCYGVPHTLEKVKER 53	
	: : : : : : : : :	
	1 KHPRRKRRYQYNSLEPISVLMSGSKMSVSKAQSITYGIPHSTLEKVKER 53	
RESULT 7		
AAE24370		
ID	AAE24370 standard; Protein; 53 AA.	
XX		
AC	AAE24370;	
XX		

04-OCT-2002 (first entry)

Fruit fly E93 programmed cell death modulating protein conserved domain.

Fruit fly; programmed cell death modulating protein; adenocarcinoma; AIDS; neurodegenerative disease; leukemia; acquired immunodeficiency syndrome; AIDS; cellular apoptosis; leukemia; acquired immunodeficiency syndrome; AIDS; neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa; Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration; aplastic anaemia; ischaemic injury; myocardial infarction; stroke; reperfusion injury; toxin-induced disease; genetic immunodeficiency; vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective; myeloma; nootropic; vasotropic; immunostimulant; cerebroprotective; cardiant; cancer; E93 protein.

Drosophila melanogaster.

WO200234882-A2.

02-MAY-2002.

29-OCT-2001; 2001WO-US48053.

27-OCT-2000; 2000US-243865P.

(UYMA-) UNIV MARYLAND BIOECHOLOGY INST.

Baehrecke EH;

WPI: 2002-479717/51.

Novel programmed cell death modulating proteins, useful for treating or preventing disorders associated with abnormal cell proliferation and apoptosis such as cancer, stroke, Parkinson's disease, myocardial infarction.

Claim 1; Fig 1; 88pp; English.

The present invention relates to novel programmed cell death modulating proteins and polynucleotides encoding such proteins. Sequences of the invention are useful to screen potential cellular apoptosis inhibiting compounds to determine their use as therapeutic agents for treatment of diseases associated with increased programmed cell death. They are also useful for treating or preventing disorders associated with decrease in apoptosis. Programmed cell death modulating sequences are useful for treating or preventing cancer e.g. adenocarcinoma, leukemia, lymphoma, melanoma, myeloma. Inhibition of the activity of the sequences of the invention are useful for treating disorders associated with increase in cell death or apoptosis such as acquired immunodeficiency syndrome (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic injuries (e.g., myocardial infarction, stroke, reperfusion injury), myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced diseases and other infectious or genetic immunodeficiencies. Sequences of the invention are used as vaccines and in gene therapy. The present sequence is fruit fly E93 programmed cell death modulating protein conserved domain.

Sequence 53 AA:

Query Match 60.4%; Score 165; DB 23; Length 53;
Best Local Similarity 60.4%; Pred. No. 4.4e-17;
Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

1 KOPRKKRGYROYDHEIMEAIAWMSGKMSVSKAGIGYVPHSTLEKYKER 53
1 KOTPRKRGYROYDHEIMEAIAWMSGKMSVSKAGIGYVPHSTLEKYKER 53

RESULT 8
ABBT1145
ID ABB71145 standard; Protein: 1140 AA.
AC ABB71145;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 40227.

Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EM;

WPI: 2001-656860/75.

N-PSDB: ABL15248.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

Disclosure; SEQ ID NO 40227; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-AB572072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 1140 AA:

Query Match 60.4%; Score 165; DB 22; Length 1140;
Best Local Similarity 60.4%; Pred. No. 2.4e-15;
Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

1 KOPRKKRGYROYDHEIMEAIAWMSGKMSVSKAGIGYVPHSTLEKYKER 53
1 KOTPRKRGYROYDHEIMEAIAWMSGKMSVSKAGIGYVPHSTLEKYKER 793

RESULT 9
AAE24372
ID AAE24372 standard; Protein: 1165 AA.
AC AAE24372;

04-OCT-2002 (first entry)

Fruit fly E93 programmed cell death modulating protein #1.

Fruit fly; programmed cell death modulating protein; adenocarcinoma; AIDS; cellular apoptosis; leukemia; acquired immunodeficiency syndrome; AIDS; neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa; Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration; aplastic anaemia; ischaemic injury; myocardial infarction; stroke; reperfusion injury; toxin-induced disease; genetic immunodeficiency; vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective; myeloma; nootropic; vasotropic; immunostimulant; cerebroprotective.

KM cardiact; cancer; E93 protein.
 XX
 OS *Drosophila melanogaster*.
 XX
 PN WO200234882-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-US48053.
 XX
 PR 27-OCT-2000; 2000US-243865P.
 XX
 XX
 PA (UWMA-) UNIV MARYLAND BIOTECHNOLOGY INST.
 XX
 PI Baehrecke EH;
 XX
 DR WPI; 2002-479717/51.
 DR N-PSDB; AAD39237.
 XX
 PT Novel programmed cell death modulating proteins, useful for treating or
 PT preventing disorders associated with abnormal cell proliferation and
 PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial
 PT interaction -
 XX
 PS Claim 9; Page 65-71; 88pp; English.

The present invention relates to novel programmed cell death modulating proteins and polynucleotides encoding such proteins. Sequences of the invention are useful to screen potential cellular apoptosis inhibiting compounds to determine their use as therapeutic agents for treatment of diseases associated with increased programmed cell death. They are also useful for treating or preventing disorders associated with decrease in apoptosis. Programmed cell death modulating sequences are useful for treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma. Inhibition of the activity of the sequences of the invention are useful for treating disorders associated with increase in cell death or apoptosis such as acquired immunodeficiency syndrome (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis pigmentosa, Parkinson's disease and cerebellar degeneration), ischemic injuries (e.g., myocardial infarction, stroke, reperfusion injury), myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced diseases and other infectious or genetic immunodeficiencies. Sequences of the invention are used as vaccines and in gene therapy. The present sequence is fruit fly E93 programmed cell death modulating protein.

Seq	Sequence	1165 AA;	Query Match	60.4%;	Score 165;	DB 23;	length 1165;
			Best Local Similarity	60.4%;	Pred. No.2.5e-15;		
	Matches	32;	Conservative	6;	Mismatches 15;	Indels 0;	Gaps 0;
Qy	1	KQPKRKGRYQYDHEIMEEAIAMVMSKMSYSKAGQIGYVPHSTLEKVER	53				
Db	758	KGLPKRRKRYNYDSDSLVEAVKRAQKREMSVHRGSGSYGVPHSTLEKVER	810				

RESULT 10	
AAE24595	
ID	AAE24595 standard; Protein: 53 AA.
XX	
AC	
XX	AAE24595;
XX	
DT	04-OCT-2002 (first entry)
XX	
DE	Nematode E93 programmed cell death modulating protein conserved domain.
XX	
KW	Nematode; programmed cell death modulating protein; adenocarcinoma;
KW	cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;
KW	neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;
KW	Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;
KW	aplastic anaemia; ischaemic injury; myocardial infarction; stroke;
KW	reperfusion injury; toxin-induced disease; genetic immunodeficiency;
KW	vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;

KW	myeloma nontropic; vasotropic;	immunostimulant; cerebroprotective;
KV	cardiant; cancer; E93 protein.	
XX		
OS	Caenorhabditis elegans.	
XX		
FN	W0200234882-A2.	
XX		
PD	02-MAY-2002.	
XX		
PF	29-OCT-2001; 2001WO-US48053.	
XX		
PR	27-OCT-2000; 2000US-243865P.	
XX		
PA	(UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.	
XX		
PI	Baehecke EH;	
XX		
DR	WPI; 2002-479717/51.	
XX		
PT	Novel programmed cell death modulating proteins, useful for treating OR	
PT	preventing disorders associated with abnormal cell proliferation and	
PT	apoptosis such as cancer, stroke, Parkinson's disease, myocardial	
PT	infarction	
XX		
PS	Claim 1; Fig 1; 88pp: English.	

The present invention relates to novel programmed cell death modulating proteins and polynucleotides encoding such proteins. Sequences of the invention are useful to screen potential cell/apoptosis inhibiting compounds to determine their use as therapeutic agents for treatment of diseases associated with increased programmed cell death. They are also useful for treating or preventing disorders associated with decrease in apoptosis. Programmed cell death modulating sequences are useful for treating or preventing cancer (e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma. Inhibition of the activity of the sequences of the invention are useful for treating disorders associated with increase in cell death or apoptosis such as acquired immunodeficiency syndrome (AIDS), neurodegenerative diseases (e.g. Alzheimer's disease, retinitis pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic injuries (e.g., myocardial infarction, stroke, reperfusion injury), myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced diseases and other infectious or genetic immunodeficiencies. Sequences of the invention are used as vaccines and in gene therapy. The present sequence is nematode E93 programmed cell death modulating protein conserved domain.

Seq	Sequence	53 AA:	Score 163:	DB 23:	Length 53:
Query Match		59.7%			
Best Local Similarity		56.6%			
Matches		30:	Mismatches 10:	Indels 0:	Gaps 0:
OY	1 KQPKKRGKRYQYDHEIMEEALIAMYSKMSVSKAQGIYGVPHSTLEKVKER	53			
Db	1 KRSRKRRQYRRYDKNALDEAVRSVRKREEMTVHRRGSEFGPHSTLEKVKER	53			

PN	WO200171042-A2.
XX	Drosophila melanogaster.
OS	
XX	
KM	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX	
DE	Drosophila melanogaster polypeptide seq ID NO 27876.
XX	
DT	26-MAR-2002 (first entry)
XX	
AC	ABB67028;
XX	
ID	ABB67028 standard; Protein; 1046 AA.
ABB67028	
RESULT 11	

XX	27-SEP-2001.
PD	
XX	
XX	23-MAR-2001; 2001WO-USO9231.
XX	
XX	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
XX	WPI; 2001-656860/75.
DR	N-PSDB; ABL11131.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
XX	interactions -
PS	
XX	Disclosure; SEQ ID NO 27876; 21pp + Sequence Listing; English.
CC	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
XQ	Sequence 1046 AA:

Query Match	32.4%	Score 88.5	DB 22,	Length 1046;
Best Local Similarity	34.7%	Pred. No. 0.0007;		
Matches 17; Conservative	14;	Mismatches 17;	Indels 1;	Gaps 1;

Oy 5KKGRYRDYDHEINEEAIAMVMSGKSMYSKAOGIYGVPSTLLEKYVKER 53
|| : : : : : || : : : : : || : : : : : || : : : : :
Db 755 KKGGGTAKWNEDALNALALELRSGOISANKASKAFGIPISTL-YKIARR 802

RESULT 12
ABB59068
ID ABB59068 standard; Protein: 1064 AA.
xx

AC ABB59068;
XX
DT 26-MAR-2002 (first entry)
XX

XX
XX
XX
KW
KV
pharmacological
developmental biology; cell signalling; insecticide;
Drosophila; Drosophila melanogaster poly(PTC)SeQ ID NO 3996.

aa
OS
xx
PN
WO200171042-A2.
Drosophila melanogaster.

xx 27-SEP-2001.
pd
xx
23-MAR-2001; 2001WO-US09231.
pf

PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX

PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI: 2001-6556860/75.

DR N-PSDB: ABL03171.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure: SEQ ID NO 3996; 21pp + Sequence Listing; English.
XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB161675) and the encoded proteins
CC (AAB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published_pat_sequences.

Sequence	1064 AA;
----------	----------

Query Match	32.4%;	Score	88.5;	DB	22;	Length	1064;
Best Local Similarity	34.7%;	Pred. No.	0.00072;				
Matches	17;	Conservative	14;	Mismatches	17;	Indels	1;
						Gaps	1;

QY 5 KKKGRVRYDDHETMEEAIAVMWGSKMSYKAQGIVGVPHSTLEKVKR 53
773 KKGGTKMWNEDALQNALLAELRLSGQISANKASKAFGISTL--YKIARR 820

RESULT	13
ABB63113	
ID	ABB63113 standard; Protein; 661 AA.

AC ABB63113;
XX
DT 26-MAR-2002 (first entry)
XX

DE DE Drosophila melanogaster polypeptide SEQ ID NO 16131.
DE XX
XX
KW Drosophila: developmental biology; cell signalling; insecticide;
KU pharmacology; cell signalling; insecticide; insecticide; insecticide;

xx Drosophila melanogaster.
os
xx W02001171042-A2
pn

AA
PD
XX
27-SEP-2001.
23-MAR-2001: 2001WO-US09231.
PF

23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX

XX
PI Venter JC, Adams M, Li PMD, Myers EW;
XX

DR	N-PSDB; ABL07216.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more

PT interactions -
XX
PS disclosure; SEQ ID NO 16131; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 15:01:44 : Search time 95 Seconds
(without alignments)
143.966 Million cell updates/sec

Title: US-10-016-768a-2

Perfect score: 273

Sequence: 1 KQPRKRRGRYRDHEIMEAIVMWSGKMSVSKAGIYGVPHSTLEYVKRER 53

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP. archaia:*
2: SP. bacteria:*
3: SP. fungi:*
4: SP. human:*
5: SP. invertebrate:*
6: SP. mammal:*
7: SP. mhc:*
8: SP. organelle:*
9: SP. phage:*
10: SP. plant:*
11: SP. rodent:*
12: SP. virus:*
13: SP. vertebrate:*
14: SP. unclassified:*
15: SP. viirus:*
16: SP. bacteriaph:*
17: SP. archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	273	100.0	213	4	Q96NK1 mus sapien
2	273	100.0	517	11	Q8CJG4 mus musculu
3	242	88.6	396	11	Q8C9Q0 mus musculu
4	242	88.6	433	11	Q8BGT2 mus musculu
5	242	88.6	572	4	Q96JN0 homo sapien
6	242	88.6	619	4	Q8N3I6 homo sapien
7	165	60.4	1165	5	Q9VD60 drosophila
8	165	60.4	1358	5	Q95YH8 apis mellif
9	163	59.7	185	5	Q22051 caenorhabd
10	88.5	32.4	645	5	Q8MKX3 drosophila
11	88.5	32.4	660	5	Q24457 drosophila
12	88.5	32.4	1064	5	Q9V5N1 drosophila
13	88.5	32.4	1085	5	Q24455 drosophila
14	78.5	28.8	652	5	Q77168 apis mellif
15	72	26.4	107	12	Q91IHO frog adenov
16	70.5	25.8	661	5	Q9V8S2 drosophila

17	69.5	25.5	1046	5	Q9W0W2 drosophila
18	69	25.3	365	3	Q9P3F3 neurospora
19	67	24.5	924	16	Q8CJL0 streptomyce
20	64	23.4	393	11	Q8C9J6 mus musculu
21	63	23.1	977	5	Q9W0K7 drosophila
22	63	23.1	977	5	Q9W0K7 drosophila
23	61	22.3	478	16	Q8R5T4 drosophila
24	61	22.3	494	16	Q8R5H9 drosophila
25	61	22.3	444	16	Q8D2T2 drosophila
26	60.5	22.2	325	3	Q9UVG7 streptococc
27	59.5	21.8	267	5	Q9GP27 magnaporthe
28	59.5	21.8	267	5	Q9V5X7 drosophila
29	59.5	21.8	551	8	Q9VF64 beta vulgar
30	59	21.6	176	2	Q8KSEA enterococcu
31	59	21.6	308	16	Q8PCX8 xanthomonas
32	59	21.6	342	16	Q8F5P1 leptospira
33	59	21.6	419	2	Q46445 chlamydia t
34	59	21.6	424	16	Q84654 chlamydia t
35	59	21.6	470	13	Q8Q6W1 anquilla ja
36	59	21.6	1067	5	Q8V0H3 drosophila
37	58.5	21.4	224	2	Q8V0Q3 staphylococ
38	58.5	21.4	224	2	Q8GFL0 staphylococ
39	58.5	21.4	224	2	Q8G8W8 staphylococ
40	58	21.2	193	4	Q8N9Y1 homo sapien
41	58	21.2	271	17	Q8TMT7 methanosarc
42	58	21.2	306	16	Q8PG12 xanthomonas
43	58	21.2	413	4	Q8IWF9 homo sapien
44	58	21.2	761	16	Q8DIA7 synecchococ
45	58	21.2	1880	5	Q18465 hirtudo med

ALIGNMENTS

RESULT 1	Q96NK1	PRELIMINARY:	PRT:	213 AA.
AC	Q96NK1			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	Hypothetical protein FLJ30696.			
OC	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,			
RA	Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,			
RA	Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,			
RA	Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,			
RA	Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,			
RA	Magatsuna M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,			
RA	Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahara K.,			
RA	Masuno Y., Nagai K., Isogai T.,			
RT	"NEO human cDNA sequencing project."			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AK055258; BAB70892.1; -			
KW	Hypothetical protein			
SO	SEQUENCE 213 AA; 23477 MW; 4D76CABF95251B2 CRC64;			

Query Match 100.0%; Score 273; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.4e-26;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	1 KQPRKRRGRYRDHEIMEAIVMWSGKMSVSKAGIYGVPHSTLEYVKRER 53
Db	124 KQPRKRRGRYRDHEIMEAIVMWSGKMSVSKAGIYGVPHSTLEYVKRER 176
RESULT 2	

```
Q8CJG4
ID 08CJG4 PRELIMINARY: PRT: 517 AA.
AC 08CJG4:
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Transcription factor MLR1.
GN MLR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain:
RA Kunieda T., Park J., Takeuchi H., Kubo T.:
RT "Mus musculus mlr1 and mlr2 mRNA for transcription factor MLR1 and
RT MUR2."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB076078; BAC20954.1; -.
SQ SEQUENCE 517 AA; 57316 MW; C97403D3D296C52E CRC64;

Query Match 100.0%; Score 273; DB 11; Length 517;
Best Local Similarity 100.0%; Pred. No. 6.4e-26;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KOPRKKRGYRQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
D 429 KOPRKKRGYRQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 481
|||||
|||||

RESULT 3
Q8C900
ID 08C900 PRELIMINARY: PRT: 396 AA.
AC 08C900:
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK041621; BAC31007.1; -.
KW Hypothetical protein.
FT NON_TER 396
SQ SEQUENCE 396 AA; 43085 MW; EE4A585FE2336E35 CRC64;

Query Match 88.6%; Score 242; DB 11; Length 396;
Best Local Similarity 86.8%; Pred. No. 3.8e-22;
Matches 46; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 KOPRKKRGYRQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
D 337 KOPRKKRGYRQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 389
|||||
|||||

RESULT 4
Q8BGT2
ID 08BGT2 PRELIMINARY: PRT: 433 AA.
AC 08BGT2:
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
```

```
DE Transcription factor MLR2 (Hypothetical protein).
GN MLR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain:
RA Kunieda T., Park J., Takeuchi H., Kubo T.:
RT "Mus musculus mlr1 and mlr2 mRNA for transcription factor MLR1 and
RT MUR2."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AB076079; BAC20955.1; -.
DR EMBL; AK041090; BAC30816.1; -.
KW Hypothetical protein.
SQ SEQUENCE 433 AA; 47124 MW; 736656D1F7E9A041 CRC64;

Query Match 88.6%; Score 242; DB 11; Length 433;
Best Local Similarity 86.8%; Pred. No. 4.2e-22;
Matches 46; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 KOPRKKRGYRQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
D 337 KOPRKKRGYRQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 389
|||||
|||||

RESULT 5
Q96JN0
ID 096JN0 PRELIMINARY: PRT: 572 AA.
AC 096JN0:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein KIAA1795 (Fragment).
GN KIAA1795.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain:
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 8:85-95(2001).
DR EMBL; AB058698; BAB47424.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 572 AA; 62730 MW; FB0A401D3F060DF4 CRC64;

Query Match 88.6%; Score 242; DB 4; Length 572;
Best Local Similarity 86.8%; Pred. No. 5.8e-22;
Matches 46; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 KOPRKKRGYRQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
D 476 KOPRKKRGYRQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 528
|||||
|||||

RESULT 6
```


[illegible]

us-10-016-768a-2.rspt

Page 7

Search completed: August 4, 2003, 15:06:47
Job time : 100 secs

THIS PAGE BLANK (USPTO)

Mon Aug 4 16:28:23 2003

us-10-016-768a-2.rapb

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 15:06:54 ; Search time 50 Seconds

(without alignments)
125.885 Million cell updates/sec

Title: US-10-016-768a-2

Perfect score: 273

Sequence: 1 KQPKRRGRYROYDHEIMEE.....KAQIGVPHSTLEYKVKR 53

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubppaa/PCIT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubppaa/PCITUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	273	100.0	53	14	US-10-016-768-2
2	273	100.0	44	14	US-10-016-768-8
3	233.5	85.5	54	14	US-10-016-768-3
4	229	83.9	53	14	US-10-016-768-4
5	165	60.4	53	14	US-10-016-768-1
6	165	60.4	1165	14	US-10-016-768-10
7	163	59.7	53	14	US-10-016-768-5
8	66	24.2	378	15	US-10-156-761-12022
9	62	22.3	325	11	US-09-938-901-2
10	61	22.3	352	9	US-09-741-659-432
11	58	21.2	364	9	US-09-886-400-4
12	58	21.2	364	14	US-10-112-357-4
13	58	21.2	364	14	US-10-114-403-4
14	58	21.2	364	14	US-10-116-606-4
15	58	21.2	364	14	US-10-112-331-4

16	58	21.2	364	14	US-10-112-377-4	Sequ
17	58	21.2	364	14	US-10-116-581-4	Sequ
18	58	21.2	364	14	US-10-112-442-4	Sequ
19	58	21.2	364	14	US-10-112-418-4	Sequ
20	58	21.2	364	14	US-10-114-083-4	Sequ
21	56.5	20.7	2027	15	US-10-234-961-2	Sequence 2, Appl1
22	55.5	20.3	104	9	US-09-815-242-11443	Sequence 11443, A
23	55.5	20.3	104	9	US-09-815-242-11597	Sequence 11597, A
24	55	20.1	204	10	US-09-738-626-5390	Sequence 5390, Ap
25	55	20.1	420	9	US-09-815-242-5343	Sequence 5343, Ap
26	55	20.1	420	9	US-09-815-242-12648	Sequence 12648, A
27	54.5	20.1	470	9	US-09-815-242-12817	Sequence 12817, A
28	54.5	20.0	420	9	US-09-815-242-13723	Sequence 13723, A
29	53.5	19.6	384	9	US-09-825-414-54	Sequence 54, Appl
30	53.5	19.6	1597	10	US-09-832-292-35	Sequence 35, Appl
31	53	19.4	102	15	US-10-156-761-12459	Sequence 12459, A
32	53	19.4	433	15	US-10-156-761-9784	Sequence 9784, Ap
33	52.5	19.2	363	9	US-09-815-242-10607	Sequence 10607, A
34	52	19.0	102	9	US-09-815-242-13196	Sequence 13196, A
35	52	19.0	102	10	US-09-925-637-38	Sequence 38, Appl
36	52	19.0	102	10	US-09-908-931B-34	Sequence 34, Appl
37	52	19.0	102	15	US-10-084-205-38	Sequence 38, Appl
38	52	19.0	129	11	US-09-764-891-2977	Sequence 2977, Ap
39	52	19.0	552	10	US-09-817-764-4	Sequence 4, Appl1
40	52	19.0	724	15	US-10-128-714-3025	Sequence 3025, Ap
41	52	19.0	1139	15	US-10-128-714-8025	Sequence 8025, Ap
42	51.5	18.9	349	11	US-09-742-153-18	Sequence 18, Appl
43	51.5	18.9	358	16	US-10-278-536-162	Sequence 162, Appl
44	51.5	18.9	534	14	US-09-742-153-22	Sequence 22, Appl
45	51.5	18.9	1399	11	US-09-742-153-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-10-016-768-2

Sequence 2, Application US/10016768

Publication No. US2002012443A1

GENERAL INFORMATION:

APPLICANT: Baehrcke, Eric H.

TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH

FILE REFERENCE: 4115-131

CURRENT APPLICATION NUMBER: US/10/016, 768

CURRENT FILING DATE: 2001-10-29

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 53

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: MISC_FEATURE

LOCATION: (1)..(54)

OTHER INFORMATION: X CAN BE ANY AMINO ACID

US-10-016-768-2

Query Match

Best Local Similarity 100.0%; Score 273; DB 14; Length 53;

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db

1 KQPKRRGRYROYDHEIMEEAIAMVSGKMSVSKAQIGVPHSTLEYKVKR 53

US-10-016-768-8

Sequence 8, Application US/10016768

Publication No. US2002012443A1

GENERAL INFORMATION:

APPLICANT: Baehrcke, Eric H.

TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH

FILE REFERENCE: 4115-131
CURRENT APPLICATION NUMBER: US/10/016,768
CURRENT FILING DATE: 2001-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 442
TYPE: PRT
ORGANISM: Homo sapiens
US-10-016-768-8

Query Match 100.0%; Score 273; DB 14; Length 442;
Best Local Similarity 100.0%; Pred. No. 1,3e-30;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
DB 353 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 405

RESULT 3

US-10-016-768-3
Sequence 3, Application US/10016768
Publication No. US20020142443A1
GENERAL INFORMATION:
APPLICANT: Baehrcke, Eric H.
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH
FILE REFERENCE: 4115-131
CURRENT APPLICATION NUMBER: US/10/016,768
CURRENT FILING DATE: 2001-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 54
TYPE: PRT
ORGANISM: T. nigroviridis
US-10-016-768-3

Query Match 85.5%; Score 233.5; DB 14; Length 54;
Best Local Similarity 81.5%; Pred. No. 1,3e-25;
Matches 44; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

OY 1 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
DB 1 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 54

RESULT 4

US-10-016-768-4
Sequence 4, Application US/10016768
Publication No. US20020142443A1
GENERAL INFORMATION:
APPLICANT: Baehrcke, Eric H.
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH
FILE REFERENCE: 4115-131
CURRENT APPLICATION NUMBER: US/10/016,768
CURRENT FILING DATE: 2001-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 53
TYPE: PRT
ORGANISM: M. musculus
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(54)
OTHER INFORMATION: X can be any amino acid
US-10-016-768-4

Query Match 83.9%; Score 229; DB 14; Length 53;
Best Local Similarity 81.1%; Pred. No. 5,6e-25;
Matches 43; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
DB 1 KHPRKKGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53

RESULT 5

US-10-016-768-1
Sequence 1, Application US/10016768
Publication No. US20020142443A1
GENERAL INFORMATION:
APPLICANT: Baehrcke, Eric H.
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH
FILE REFERENCE: 4115-131
CURRENT APPLICATION NUMBER: US/10/016,768
CURRENT FILING DATE: 2001-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 53
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(54)
OTHER INFORMATION: X can be any amino acid
US-10-016-768-1

Query Match 60.4%; Score 165; DB 14; Length 53;
Best Local Similarity 60.4%; Pred. No. 5,6e-16;
Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

OY 1 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
DB 1 KGRPKRGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53

RESULT 6

US-10-016-768-10
Sequence 10, Application US/10016768
Publication No. US20020142443A1
GENERAL INFORMATION:
APPLICANT: Baehrcke, Eric H.
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH
FILE REFERENCE: 4115-131
CURRENT APPLICATION NUMBER: US/10/016,768
CURRENT FILING DATE: 2001-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 1165
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-10-016-768-10

Query Match 60.4%; Score 165; DB 14; Length 1165;
Best Local Similarity 60.4%; Pred. No. 2,7e-14;
Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

OY 1 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 53
DB 758 KGRPKRGRYQYDHEIMEAIAVMGSKMSVSKAOGIYGVPHSTLEYKVKR 810

RESULT 7

US-10-016-768-5
Sequence 5, Application US/10016768
Publication No. US20020142443A1
GENERAL INFORMATION:
APPLICANT: Baehrcke, Eric H.
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH
FILE REFERENCE: 4115-131
CURRENT APPLICATION NUMBER: US/10/016,768
CURRENT FILING DATE: 2001-10-29

NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 53
TYPE: PRT
ORGANISM: Caenorhabditis elegans
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)-(54)
OTHER INFORMATION: X CAN BE ANY AMINO ACID
US-10-016-768-5

Query Match
Best Local Similarity 59.7%; Score 163; DB 14; Length 53;
Matches 30; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 KOPRRKRRYROYDHEIMEAIAWMSGKMSVSKAGIYVPHSTLEYKVR 53
DB 1 KRSRPRKRGYRKYDKNALDEAVRSVRGEMTVHRAGSFVPHSTLEYKVR 53

RESULT 8
US-10-156-761-12022
Sequence 12022, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 12022
LENGTH: 378
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-12022

Query Match
Best Local Similarity 24.2%; Score 66; DB 15; Length 378;
Matches 12; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 3 PRKRGYROYDHEIMEAIAWMSGKMSVSKA 35
DB 58 PEKENSRYEKFDYPIKEQVANLTKGKGVAYA 90

RESULT 9
US-09-938-901-2
Sequence 2, Application US/09938901
Publication No. US20030008291A1
GENERAL INFORMATION:
APPLICANT: Kuramitsu Seiki,
APPLICANT: Yokoyama Shigeyuki
TITLE OF INVENTION: GENE ENCODING DNA REPAIR ENZYME
FILE REFERENCE: PH-1261-US
CURRENT APPLICATION NUMBER: US/09/938,901
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: JP2001-47762
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2

LENGTH: 325
TYPE: PRT
ORGANISM: Thermus thermophilus
US-09-938-901-2

Query Match
Best Local Similarity 22.7%; Score 62; DB 11; Length 325;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 2;

QY 3 PRKRGYROYDHEIMEAIAWMSGKMSV-----SKAGIYV 42
DB 208 PRKRRAR-----EERLVALVLLGKRGVLERLEGRFGLYGV 245

RESULT 10
US-09-741-669-432
Sequence 432, Application US/09741669
Patent No. US20020022718A1
GENERAL INFORMATION:
APPLICANT: Forsyth, R. Allyn
APPLICANT: Ohlsen, Kari L.
TITLE OF INVENTION: Genes identified as required for
FILE REFERENCE: ELITRA.009A
CURRENT APPLICATION NUMBER: US/09/741,669
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 432
LENGTH: 352
TYPE: PRT
ORGANISM: Escherichia coli
US-09-741-669-432

Query Match
Best Local Similarity 22.3%; Score 61; DB 9; Length 352;
Matches 17; Conservative 15; Mismatches 15; Indels 12; Gaps 3;

QY 5 KRGYROYDHEI-----MEEAIAWMSGKMSVSKAGIYVPH-STLEYKVR 53
DB 271 KRGYRAIRYDFQVDAYKCDLKEIYSLVGNKINNGHSKIT--KHFNDLEKIEER 327

RESULT 11
US-09-886-400-4
Sequence 4, Application US/09886400
Patent No. US20020045226A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE
FILE REFERENCE: DIVER110-4
CURRENT APPLICATION NUMBER: US/09/886,400
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 364
TYPE: PRT
ORGANISM: Thermococcus alcaliphilus
US-09-886-400-4

Query Match
21.2%; Score 58; DB 9; Length 364;

Best Local Similarity 36.4%; Pred. No. 7;
Matches 16; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

OY 1 KOPRKRGRYROYDHEIME--EATAMVSGMSYSKAQGIYGP 42

Db 165 KAQREKRFRYISYLLGLRELKRAIKIVFEGKVTLKAVKDIEAVP 208

RESULT 12

US-10-112-357-4

; Sequence 4, Application US/10112357
; Publication No. US20020115099A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE THEREOF
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112.357
; PRIOR APPLICATION NUMBER: 2002-03-29
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-357-4

Query Match

Best Local Similarity 36.4%; Pred. No. 7;
Matches 16; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

OY 1 KOPRKRGRYROYDHEIME--EATAMVSGMSYSKAQGIYGP 42

Db 165 KAQREKRFRYISYLLGLRELKRAIKIVFEGKVTLKAVKDIEAVP 208

RESULT 13

US-10-114-403-4

; Sequence 4, Application US/10114403
; Publication No. US20020115100A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE THEREOF
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/114.403
; PRIOR APPLICATION NUMBER: 2002-04-01
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-114-403-4

Query Match 21.2%; Score 58; DB 14; Length 364;

Best Local Similarity 36.4%; Pred. No. 7;
Matches 16; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

OY 1 KOPRKRGRYROYDHEIME--EATAMVSGMSYSKAQGIYGP 42

Db 165 KAQREKRFRYISYLLGLRELKRAIKIVFEGKVTLKAVKDIEAVP 208

RESULT 14

US-10-116-606-4

; Sequence 4, Application US/10116606
; Publication No. US20020119515A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE THEREOF
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/116.606
; PRIOR APPLICATION NUMBER: 2002-04-03
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-116-606-4

Query Match 21.2%; Score 58; DB 14; Length 364;

Best Local Similarity 36.4%; Pred. No. 7;
Matches 16; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

OY 1 KOPRKRGRYROYDHEIME--EATAMVSGMSYSKAQGIYGP 42

Db 165 KAQREKRFRYISYLLGLRELKRAIKIVFEGKVTLKAVKDIEAVP 208

RESULT 15

US-10-112-331-4

; Sequence 4, Application US/10112331
; Publication No. US20020119550A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE THEREOF
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112.331
; PRIOR APPLICATION NUMBER: 2002-03-29
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-331-4

US-10-112-331-4

Query Match	21.2%;	Score 58;	DB 14;	Length 364;
Best local similarity	36.4%;	Pred No 7;		

Best Match	Best Local	Similarity	Matches	Conservative	Mismatches	Indels	Gaps
32.14%	36.4%	32.14%	16	8	18	2	1

Matches 16; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

QY 1 KQPRKKRGYRQYDHEIME--EALAMVMSGKMSVSKAQGIYGP 42

Db 165 KAQREKRRYISYLLGLRELRAIKLVFEQVTLKAAVKDIEAVP 208

Search completed: August 4, 2003, 15:15:50
Job time : 51 secs

Job time : 51 secs

THIS PAGE BLANK



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 100337

TO: Minh-Tam Davis
Location: CM-1/8A01/8E12
Art Unit: 1642
Monday, August 04, 2003

Case Serial Number: 10/016768

From: Toby Port
Location: Biotech-Chem Library
CM1-6A04
Phone: 308-3534

toby.port@uspto.gov

Search Notes

Dear Examiner Davis,

Here are the results of your search.

Please feel free to contact me if you have any questions.

Toby Port

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK